PLASMONIC DEVICES IN THE TERAHERTZ AND OPTICAL FREQUENCY DOMAINS

A Dissertation

by

HASAN TAHIR ABBAS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Robert D. Nevels
Committee Members,	Krzysztof A. Michalski
	Chin B. Su
	M. Suhail Zubairy
Head of Department,	Miroslav M. Begovic

August 2017

Major Subject: Electrical Engineering

ABSTRACT

We are living in an age where the evolution of semiconductor devices and components is contingent upon their miniaturization and seamless integration with the rest of the circuitry. Unfortunately, it is anticipated that electronic systems will soon approach the theoretical design limits of size and bandwidth and it poses to be a serious concern for the development of high-speed information technologies. Replacement of electronic pulses that act as communication signals with electromagnetic surface waves offers a very promising solution particularly in terms of device miniaturization and the heart of this optimism are the plasmonic waves arising due to collective electron oscillations at the surface of a conductor.

Surface plasmon polaritons propagating along a metal-dielectric interface at optical frequencies have lately been a subject of immense research interest mainly due to their reduced wavelength at least by an order of magnitude. Hence, miniaturized wave devices can be created at optical frequencies. Terahertz plasma waves, on the other hand, exist in infinitesimally thin plasma regions formed inside a transistor substrate, and are observed at much lower frequencies in the far-infrared regime. Due to essentially a two-dimensional nature of the plasma region, a much higher wavelength reduction factor that can exceed well beyond 100 is achievable. Furthermore, the boundary conditions due to the transistor terminals along with electric biasing create unstable resonance conditions that eventually lead to radiation in the terahertz frequency range. Such phenomena provide bright prospects for creating highly miniaturized terahertz devices.

A reliable and efficient electromagnetic (EM) analysis for multilayer geome-

tries has gained further significance due to the emergence of plasmonic structures in the optical as well as terahertz frequency domains. In this regard, integral equation (IE) techniques are ideally suited due to their efficient handling of multilayer structures. Although the presence of thin layers poses a challenge to any EM analysis technique, here the procedure is simplified due to the infinitesimally thin nature of the plasma region, which can be analyzed as a conducting sheet, with the same current flowing on either side of the sheet. Essential to any IE technique is an efficient and systematic formulation of Green functions (GFs) and their subsequent computation. In this dissertation, a transmission-line network based approach is adopted to derive spectral domain GFs for an infinitesimally thin sheet in a layered medium. The associated spatial domain counterparts are then computed through the Sommerfeld integrals (SIs).

The extraordinary electromagnetic properties of plasmonic devices are demonstrated by a presentation of the properties of plasmonic antennas and a super-resolution imaging scheme which is able to resolve objects separated only by a few nanometers.

DEDICATION

To Ammi, Abu, Tooba and Ayesha.

ACKNOWLEDGMENTS

All praise be to Allah, the most gracious and the most merciful.

I am most fortunate and blessed to have Dr. Robert Nevels as my advisor who made himself available to me, gave me excellent technical advice, provided me freedom to pursue personal research and academic interests and above all taught me to become a better human being. I feel privileged to learn from Dr. Krzysztof Michalski who helped me immensely to analyze layered media through his papers and answering my questions. I would like to thank Dr. Chin Su, who introduced me to plasmonics through the nanophotonics. Being a Pakistani, it has been an honor to work and collaborate with Dr. Suhail Zubairy who suggested on investigating imaging aspects through the surface wave structure discussed in this dissertation.

I would like to thank my colleague and friend Jongchul Shin, with whom I shared my office, for all the philosophical discussions and in particular making my stay pleasurable. I am grateful to Muhammad Al-Khalidi, for providing the motivation to get things done in a timely manner. My former colleagues, Sean Goldberger, Justin Erdle, Sungyun Jun for helping me settle in the Electromagnetics and Microwave lab (EML). I must also thank Feyza Berber for clarifying various practical aspects of semiconductor fabrication specific to this work.

I am indebted to all the great teachers starting from kindergarten to my time at Texas A&M University. I especially thank my high-school chemistry teacher Rana Muhammad Akram whose mentoring ensured that I remain focused in achieving educational goals. I owe a lot to the Islamic community in College Station that was always there for moral and spiritual support. A special thanks

goes to the Khan family for making me part of the family and taking care of my family. Thanks goes to my friends Sabeeh Ahmed, Sameer Farooqi, Umer Misgar, Bilal Wajid, Shayan Arshed and Yusuf Dogan for making my time in College Station a fond memory. I thank Hammad Akram, Ali Yazdani, Salim Butt and Ahmad Saad who stayed in touch with me and frequently inquired about my progress and well-being.

I am grateful to my mother Yasmeen Tahir and my father, Tahir Abbas for their utmost love, support, understanding and prayers. I thank my maternal aunt Nasreen Safdar for her love towards me during my time in Lahore, my brother Saad and sister Rabea for the support and encouragement. Finally, I would like to thank my dear wife, Tooba Ather who patiently endured so much during our time in College Station, from taking care of our daughter Ayesha all by herself, to letting me focus on this work. Without her, this dissertation would not have become a reality.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Robert D. Nevels [advisor], Krzysztof A. Michalski and Chin B. Su of the Department of Electrical and Computer Engineering and Professor M. Suhail Zubairy of the Department of Physics.

All work for the dissertation was completed by the student, in collaboration with Xiaodong Zeng of the Department of Physics.

Funding Sources

Graduate study was supported by a fellowship from the Fulbright Program.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

In present times, the omni-presence of telecommunication gadgetry in our lives has necessitated the need for device miniaturization without any compromise on the performance. We are witnessing an age in which microwave frequency based communication systems, that have been at the forefront of the wireless revolution for the past three decades, are reaching their saturation in terms of performance. Currently, millimeter wave devices are emerging that are the driving force of the so-called *5G movement*. Following the trends of telecommunication, it is predicted that terahertz frequency based communication systems will be ubiquitous in the near future.

Until recently, terahertz (THz) frequency systems had been overlooked compared to optical and microwave counterparts, mainly due to a lack of compact terahertz sources with sufficient power for wireless applications. Today, with the technological advancement in the semiconductor fabrication, the terahertz field has grown exponentially. It is now possible to engineer terahertz sources as well as detectors from devices that are derived from conventional field-effect transistors [3, 4, 5]. By epitaxially growing a few nanometers thick layers of group III-V semiconductors, THz quantum-cascade lasers have been realized [6]. Terahertz emitters that can be tuned are fabricated utilizing the instability in an electrically-driven plasma [7] that exists in the form of a thin sheet of free charges known as a two-dimensional electron gas (2DEG). Prediction of spontaneous emission in the terahertz frequency regime was made [3], in light of a current passing through the 2DEG.

The plasma wave propagation mechanism due to the current driven instabil-

ities in the 2DEG is analogous to surface plasmon polariton (SPP), which is a type of surface wave existing at a metal-dielectric interface at optical frequencies. Lately, there has been an immense interest in graphene material and its prospects in the development of terahertz frequency sources and detectors. While, the electronic properties of graphene are unparalleled, its integration into electronic systems has proved to be very difficult thus far. On the other hand, 2DEG based devices that are fabricated out of a III-V semiconductor heterostructure can be easily integrated with silicon based electronic devices.

Electromagnetic analysis plays a pivotal role in designing energy efficient and high performance communication systems to which antennas, backed by particular types of transmission line (TL) networks, serve as the front-end of transmitting and receiving modules. The design process of microwave systems involves an extensive use of commercial simulation tools that are mostly based on finite element method (FEM) or finite difference time domain (FDTD) techniques. Unfortunately, such tools turn out to be extremely inefficient when the thickness of the simulated object is much smaller than the wavelength of interest. Integral equation (IE) techniques employing method of moments (MoM) are ideally suited to analyze infinitesimally thin objects. Essential to any IE/MoM based technique is the formulation and the subsequent computation of Green functions (GFs) associated with the structure under observation. A 2DEG based terahertz system involves multiple layers in which an infinitesimally thin sheet described by surface conductivity is embedded. The GFs for a transistor structure in which an infinitesimally thin 2DEG layer is embedded, can be formulated following a TL network approach [8]. The fields are then extracted from the GFs via Sommerfeld Integral (SI) analysis.

The existence of an infinitesimally thin plasma region can be considered as

a realization of 2D materials that display many interesting physical properties, chief among them is the subwavelength surface wave propagation. As a result, the physical dimensions at which a structure resonates, becomes much smaller than the corresponding freespace wavelength. Currently, antennas incorporated on a semiconductor chip occupy a large amount of space. With the plasma based technology which is discussed, miniaturized antennas that can be integrated on to the semiconductor chip can be actualized.

1.1 Outline

In this section, the structure of this dissertation is briefly summarized.

Surface wave propagation was first observed in the optical frequency region where surface plasmon polaritons (SPPs) propagate along the interface of a metal, such as gold, and a dielectric. Section 2 reviews in detail the physical conditions required for exciting SPPs on the interface in light of the material properties of metals. Applications of the subwavelength properties of SPPs are discussed by studying various nanostructures as well as the design criteria that form the basis of optical antennas.

Section 3 presents the derivation procedures of spectral domain GFs using the TL-GF approach [8, 2] and extends it to incorporate infinitesimally thin sheets. The spatial domain GFs for vector potentials are obtained through the Sommerfeld integrals using the mixed potential integral equation (MPIE) method.

A surface integral equation scheme is presented in Section 4 that is exact for infinitesimally thin layers such as the ones found in a semiconductor heterostructure. The obtained results are compared with other techniques that are based on approximate boundary conditions.

The dispersion relation and its visual representation in the form of a disper-

sion diagram characterize the subwavelength properties of the 2D plasma waves. In Section 5, the dispersion diagrams of semiconductor heterostructures with 2DEG embedded, are numerically computed using a complex-valued root search algorithm known as the derivate-free argument principle method (APM).

A super-resolution, nanoscale imaging scheme is presented in Section 6 that demonstrates the subwavelength imaging capabilities of a 2DEG based system in the terahertz frequency region.

Section 7 provides concluding remarks and recommends future research on the subjects of plasmonic structures in the optical and terahertz frequency domains.

2. OPTICAL NANOANTENNAS*

An overview of the field of optical plasmonic antennas is presented in this section. After a brief introduction and historical review, the theory of surface plasmon polaritons which leads to a set of overall observations as to the requirements and restrictions placed on the operation of plasmonic waveguides and antennas is presented. Both a single metal-dielectric interface and two interfaces between a metal sheet with dielectrics on either side are considered. In the second section, the physical principles of operation and mathematical design criteria are presented for several common optical antennas including on-surface metallic structures and free-standing particles. The third section covers the basic theory of aperture radiators along with a more detailed description of some popular designs. Current applications of optical nanoantennas are presented along with a discussion on future directions in the optical nanoantenna research.

2.1 Introduction

Optical nanoantennas (or simply antennas) are nano-sized objects that transmit or receive electromagnetic fields through their intrinsic plasmonic behavior in the optical or near infrared frequency range. Optical antennas have become the object of great interest to engineers due to rapid advancements in nano-device fabrication technologies. In particular, commercial computer aided design (CAD) programs that allow the characterization of materials with a negative real part permittivity and the lower cost and advancements in electron beam (e-beam) etching machines, permitting the creation of nano-circuits with line widths on the or-

*Reprinted with permission from Optical Nanoantennas *Handbook of Antenna Technologies*, Nevels, R. D. & Abbas, H. T., 2016, Springer Singapore. Copyright 2016 Springer Science+Business Media Singapore.

der of 10 nm or smaller, have made it possible for universities and research labs to not only numerically simulate, but also build models and measure the properties of optical antennas. The fact that practical research and development is now available to the academic and industrial community has led to several new applications and technologies including dramatically improved spectroscopy [9, 10, 11], disease and toxin sensors [12, 13], wireless communication with nano-circuitry [14], and the creation of nano-circuits using subwavelength lithography [15, 16]. The opportunities for antenna engineers in this vast new landscape are great, but much work remains to be carried out on analytical and numerical characterization as well as in developing an understanding as to how electromagnetic and quantum processes can be incorporated into a coherent mathematical theory that will allow rapid and systematic design of such devices.

Dimensions of optical antennas are in the nanometer (nm) range of visible and near infrared light. Typical antenna element shapes include the dipole, bowtie, and aperture as shown in Fig. 2.1. Optical antennas are always fabricated out of noble metals, most commonly gold and silver and less commonly with aluminum, chromium, and copper. Unfortunately, there are some negative aspects of these materials that make them difficult to work with in engineering applications. In an ambient environment silver can form a silver sulfide layer that inhibits the propagation of plasmon waves [17] and as such, is not always a good candidate for an optical antenna. Aluminum has a higher dielectric function imaginary part than does gold, increasing from a small difference at 550 nm to a significantly larger value at 830 nm, where it becomes exceptionally lossy. During the fabrication of a nanoantenna, ion-beam bombardment can cause gold to melt making it hard to form a smooth metallic structure that is true to its intended design [18]. Nevertheless, noble metals are necessary for the applications

described in this section because their plasma frequency, a crucial element of the design, lies in the range of visible light. A majority of current research papers

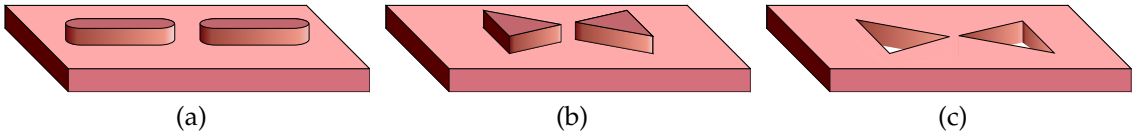


Figure 2.1: Common nanoantenna elements (a) dipole (b) bowtie and (c) aperture. The dipole, bowtie, and aperture antennas are typically etched on a low-loss substrate such as silicon dioxide.

focus on free-space designs, but for most practical applications, nanoantennas will be etched on a substrate. The type of substrate depends on the particular application, to some extent determined by the mechanism used to excite the antenna, but for many industrial applications, cost and availability considerations have driven the move to silicon dioxide SiO_2, often described in technical papers simply as 'glass'. Typical signal generation is via off-board laser light or by an on-board transmission line. Unlike classical antenna signals which can be generated at virtually any microwave carrier frequency, there are a select few frequencies where low-cost optical sources are available, although specialized lasers can operate at frequencies throughout the optical spectrum. In optical terminology the signal has traditionally been specified in wavelengths rather than frequency, however here both terms will be used. Some common wavelengths where research is being carried out are in the neighborhood of 550 , 630 , and 820 nm. An extensive list of various laser types along with their operating frequencies can be found [19].

A counter-intuitive oddity of noble metals at optical frequencies is that they

are no longer perfect conductors, but rather they have some of the properties of a dielectric with, as mentioned above, a permittivity whose real part is negative. This turns out to be a remarkable advantage over dielectrics with a positive real part, due to the fact that the wavelength of the current in these metals can be much less than the wavelength of the free-space incident field. A solid nanorod still resonates around one-half wavelength of its surface current but, given an incident field frequency, one must use numerical or approximate analytical methods to determine the shorter than free-space wavelength of the antenna at resonance. While optical antennas make excellent resonators, due in part to a tightly bound plasmon current, the wavelength mismatch between air and metal, which does not exist at microwave frequencies, is a significant factor hampering their radiation and receiving properties. To understand how the radiation efficiency of an optical antenna can be improved, it is necessary to first study the unusual properties of waves on metals at optical frequencies. The essentials of optical wave propagation and radiation are presented in the following section.

2.2 Electromagnetic Theory of Surface Plasmon Polaritons

The physical mechanism for wave propagation at optical frequencies is very different from that of propagating waves at microwave frequencies although the mathematics for determining the wave behavior is virtually the same and can be laid out in a classical Sommerfeld integral analysis. At optical frequencies surface plasmon polaritons (SPPs), electromagnetic surface waves created by coherent charge oscillations in an electron gas in a metal, propagate at a metal-dielectric interface [20, 21, 22]. Resonant plasmonic oscillations can also occur in the confined space of nanoparticles [10]. Fortunately, in most cases where optical antennas are concerned, the quantum mechanisms which create SPPs or local

confined plasmons can be expressed in terms of the material properties of metals and nanoparticles, such as shape, size and permittivity [23]. Because the study of nanoantennas relies on the behavior of plasmonic waves, it resides in the research area described as 'plasmonics', a subfield of nanophotonics [24, 25]. In this section, some of the basic theory of SPPs will be presented to aid the reader in understanding the underlying principles that determine the behavior of nanoantennas. First, the fundamental problem of electromagnetic propagation on a planar dielectric-metal interface at optical frequencies is considered, followed by a discussion on the role played by quantum effects in determining the allowable frequency range and loss mechanisms in nano-device design. This presentation does not require an understanding of quantum mechanics, but occasionally the quantum electronics terminologies are used to describe the effects that determine the properties of the constitutive parameters at optical frequencies.

2.2.1 Single Boundary Structure

Assuming a metal can have a complex permittivity at optical wavelengths, the propagation constants k_{xi} and k_{zi} at the planar interface between dielectric and metal half-spaces, $i = 1, 2$ respectively, can be found starting with expressions for transverse magnetic (TM) plane waves in these two regions as shown in Fig. 2.2. The electromagnetic field expressions in region 1 where $z \geq 0$ are,

$$\mathbf{E}_1 = \left(\hat{\mathbf{x}} E_{x1} + \hat{\mathbf{z}} E_{z1}\right) \mathrm{e}^{-j(k_x x + k_{z1} z)}, \tag{2.1a}$$

$$\mathbf{H}_1 = \hat{\mathbf{y}} H_{y1} \, \mathrm{e}^{-j(k_x x + k_{z1} z)} \tag{2.1b}$$

and in region 2 where $z \leq 0$,

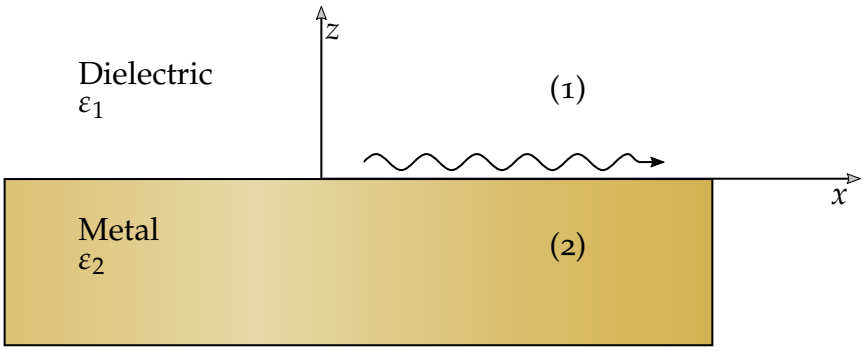


Figure 2.2: Dielectric and metal half-spaces with a planar boundary.

$$\mathbf{E}_2 = (\hat{\mathbf{x}} E_{x2} + \hat{\mathbf{z}} E_{z2})\, \mathrm{e}^{-j(k_x x - k_{z2} z)}, \tag{2.2a}$$

$$\mathbf{H}_2 = \hat{\mathbf{y}} H_{y2}\, \mathrm{e}^{-j(k_x x - k_{z1} z)}. \tag{2.2b}$$

A positive sign is chosen for k_{z2} to ensure propagation in the negative z direction. The $\mathrm{Im}(k_z)$ must therefore be negative in order to be bounded at infinity.

Ampere's law, $\nabla \times \mathbf{H} = -j\omega\varepsilon\, \mathbf{E}$, applied to (2.1) and (2.2) yields the boundary conditions:

$$k_{z1} H_{y1} = \omega \varepsilon_1 E_{x1} \tag{2.3a}$$

$$k_{z2} H_{y2} = -\omega \varepsilon_2 E_{x2} \tag{2.3b}$$

where $\varepsilon_{1,2}$ are the permittivities of the dielectric and metal regions respectively. Since metal at optical frequencies has the behavior of a lossy dielectric, continuity of tangential electric and magnetic fields, $E_{x1} = E_{x2}$ and $H_{y1} = H_{y2}$, can be applied at the boundary along with (2.3). This gives the dispersion relation for waves on the metal-air boundary:

$$\frac{k_{z1}}{\varepsilon_1} + \frac{k_{z2}}{\varepsilon_2} = 0. \tag{2.4}$$

Helmholtz equation, $\nabla^2\mathbf{E} + k_i^2\mathbf{E} = 0$, where $i = 1, 2$, and assuming the permeabilities of all regions are that of air, lead to the dispersion equations for the two regions,

$$k_x^2 + k_{zi}^2 = \varepsilon_i \left(\frac{\omega}{c}\right)^2 = \varepsilon_i k_0^2 \tag{2.5}$$

where c is the speed of light in air. Equations (2.4) and (2.5) are combined to produce the SPP propagation constant:

$$k_x = k_0 \sqrt{\frac{\varepsilon_{r1}\varepsilon_{r2}}{\varepsilon_{r1} + \varepsilon_{r2}}} \tag{2.6}$$

where k_0 is the free-space wavenumber and the subscript r is used to indicate a dielectric constant quantity. If the permittivity of the dielectric is real and that of the metal is complex, $\varepsilon_2 = \varepsilon_2' - j\varepsilon_2''$, and $\left|\varepsilon_2'\right| \gg \varepsilon_2''$ the complex propagation constant in (2.6) can be expressed as [22]:

$$k_x = k_0 \left(\frac{\varepsilon_1\varepsilon_2'}{\varepsilon_1 + \varepsilon_2'}\right)^{1/2} - jk_0 \frac{\varepsilon_2''}{2\,\varepsilon_2'^2} \left(\frac{\varepsilon_1\varepsilon_2'}{\varepsilon_1 + \varepsilon_2'}\right)^{3/2} = k_x' - jk_x''. \tag{2.7}$$

Similarly, from (2.5) and (2.6), k_{zi} becomes approximately,

$$k_{z1} = k_0 \left(\frac{\varepsilon_1^2}{\varepsilon_1 + \varepsilon_2'}\right)^{1/2} + jk_0 \frac{\varepsilon_1\varepsilon_2''}{2\left(\varepsilon_1 + \varepsilon_2'\right)^{3/2}} = k_{z1}' + jk_{z1}'', \tag{2.8a}$$

$$k_{z2} = k_0 \left(\frac{\varepsilon_2'^2}{\varepsilon_1 + \varepsilon_2'}\right)^{1/2} - jk_0 \frac{\varepsilon_2''\left(2\varepsilon_1 + \varepsilon_2'\right)}{2\left(\varepsilon_1 + \varepsilon_2'\right)^{3/2}} = k_{z2}' - jk_{z2}''. \tag{2.8b}$$

From (2.4)-(2.8), some very important conclusions can be drawn regarding the nature of the two half-space materials and the properties of nano surface waves on their intersecting boundary:

1. First, if loss is neglected ($\varepsilon_2'' \simeq 0$), a valid assumption for noble metals over

significant segments of the optical range so that the permittivities of both regions are real, the equality in (2.4) will not hold if these permittivities are both positive or both negative. But, if the permittivity in the dielectric region 1 is positive, $\varepsilon_1 > 0$, and the real part of the permittivity in the metal region 2 is negative, $\varepsilon_2' < 0$, then it is possible to satisfy (2.4).

2. Nevertheless, a wave still cannot exist if the term under the square root in (2.6) is negative because this would lead to a totally complex k_x, which would mean the fields in (2.1) and (2.2) will exponentially grow or decay rather than propagate along the boundary. However, if

$$\varepsilon_1 > 0, \quad \varepsilon_2' < 0, \quad \text{and} \quad \left|\varepsilon_2'\right| > \varepsilon_1, \tag{2.9}$$

then the term under the square root in (2.6) is positive and propagation of a wave at the boundary is assured. This condition (2.9) causes the signs of the terms under the square roots in (2.8) to be negative, but (2.4) will still hold as long as same sign for $\sqrt{-1}$ is chosen for both square roots.

3. If the conditions in (2.9) are met then from (2.6), $k_x > k_0$ and therefore k_z in (2.5) must be complex. According to (2.1) and (2.2), this results in

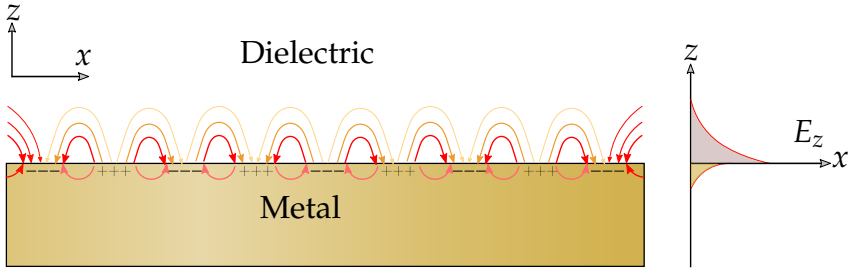


Figure 2.3: Surface plasmon propagation along a dielectric-metal boundary and exponential decay perpendicular to the boundary.

exponential decay of the field away from the boundary in both the positive and negative z-directions as long as $\sqrt{-1} = -j$ is chosen for the square root terms in (2.8). Exponential decay of the plasmon surface wave into the dielectric and metal is illustrated in Fig. 2.3.

4. According to Snell's law, the component of the wavenumber k_x tangential to the boundary is the same in both mediums. If medium 1 is air, the x-component of a plane wave incident upon the boundary at an angle θ from normal is $k_{x1} = k_0 \sin\theta$, which is less than k_0. However, according to (2.6) and (2.9) surface plasmons cannot exist unless $k_x > k_0$. Therefore, surface plasmons cannot be excited by shining light on a flat metal surface.

5. From (2.7), the velocity of a surface plasmon wave on a planar boundary is:

$$v_{sp} = \frac{\omega}{k'_x} = c\left(\frac{\varepsilon_1 + \varepsilon'_2}{\varepsilon_1 \varepsilon'_2}\right)^{1/2} \tag{2.10}$$

and its wavelength is:

$$\lambda_{sp} = \frac{2\pi}{k'_x} = \left(\frac{\varepsilon_1 + \varepsilon'_2}{\varepsilon_1 \varepsilon'_2}\right)^{1/2} \tag{2.11}$$

6. The exponential decay of a plasmon along its propagation direction is determined by the second term in (2.7). The 'propagation length' of a plasmon wave is the distance $x = L$ at which the wave decays to $1/e$ of its initial value. The plasmon propagation length on a planar surface is therefore:

$$L = \frac{1}{k'_x} = \frac{2\,\varepsilon'^2_2}{k_0\,\varepsilon''_2}\left(\frac{\varepsilon_1 + \varepsilon'_2}{\varepsilon_1 \varepsilon'_2}\right)^{1/2} \tag{2.12}$$

7. If the conditions in (2.9) hold, the second term on the right hand side in (2.8a) and (2.8b) are small so, neglecting these terms, perpendicular to the boundary the $1/e$ decay distance into the dielectric is [22]:

$$z_1 = \left| \frac{1}{k'_{z1}} \right| = \frac{1}{k_0} \left| \frac{2\,(\varepsilon'_2)^2}{k_0\,\varepsilon''_2} \left(\frac{\varepsilon_1 + \varepsilon'_2}{\varepsilon_1^2} \right)^{1/2} \right| \tag{2.13}$$

and,

$$z_2 = \left| \frac{1}{k'_{z2}} \right| = \frac{1}{k_0} \left| \frac{2\,(\varepsilon'_2)^2}{k_0\,\varepsilon''_2} \left(\frac{\varepsilon_1 + \varepsilon'_2}{\varepsilon_2'^2} \right)^{1/2} \right| \tag{2.14}$$

into the metal.

8. Power and energy relationships can be found by assuming a value for the magnetic field amplitude and by using Ampere's law to determine the electric field.

The conditions described in (2.9) are met by noble metals in air for the optical and near-infrared ranges in the case of a TM polarized wave. A similar analysis will show that a transverse electric (TE) polarized wave cannot exist at a dielectric-metal boundary.

In practical applications, one should take care in the selection of the metal to be used in a nanoantenna design. Some metals, such as silver, tend to develop an oxidation layer, which can severely decrease propagation length and increase the antenna impedance. For a more detailed discussion concerning the effect of a silver sulfide oxidation layer on a silver substrate plasmon wave, the reader is referred to [17]. Fig. 2.4 shows the propagation length as a function of wavelength for three metals, gold, silver and aluminum [26]. The practitioner is cautioned that although the propagation length of a surface plasmon does increase as fre-

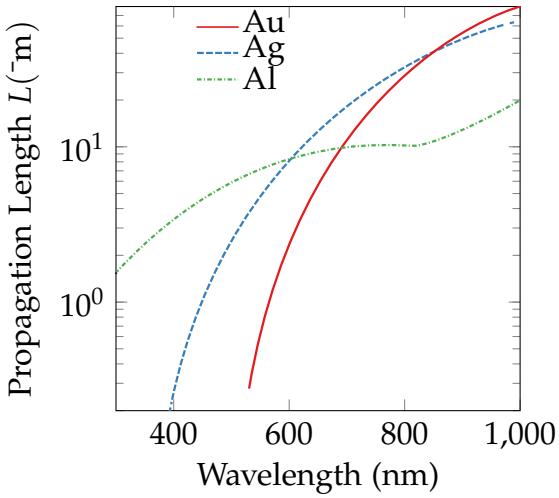


Figure 2.4: Propagation length of a surface plasmon propagating along the interface between a dielectric (refractive index 1.32) and a metal as a function of wavelength calculated for gold (Au), silver (Ag), and aluminum (Al).

quency decreases, the analysis leading to this figure is based on an idealization of the physical problem. Here, a pure wave propagation condition was assumed in the formulation of equations (2.1) and (2.2). In the practical case, a source such as the feed point of an antenna, must be included in the problem formulation. If the source is included, one sees that as the frequency decreases, less energy is transferred into the surface plasmon wave and more into the space wave; the field that is radiated directly into space from the source point. In the far infrared frequency range, the surface plasmon essentially ceases to exist. Plasmon wave behavior is discussed in more detail in the section of this section concerning aperture antennas.

2.3 Quantum Effects

The above analysis considers only the classical electromagnetic aspects of propagation on a dielectric metal boundary. However, quantum mechanical ef-

fects inherent in metals and dielectrics enter the picture through the permittivity, thereby determining its value as well as the frequency range over which SPPs can propagate. Essentially, quantum processes reveal the frequency range over which the metal has a permittivity with a negative real part and where losses begin to have a serious effect on wave propagation within that frequency range. In the following, the influence of plasma resonance and atomic collisions on SPP propagation are briefly explored. Although the properties described below are best understood using quantum mechanics, the mathematics below is classical, primarily developed by concatenating harmonic oscillator models. However, the differences in the results of the two theories is small, allowing classical electromagnetic calculations without the necessity of delving into quantum theory. Free electrons can exist near the top of the energy band distribution in a metal. The interaction of these electrons with photons and the long-range Coulomb force of atoms create an electron oscillation known as a plasmon. Taking these effects into account, the dielectric function of metal can be expressed in terms of the Drude model [27]:

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega - j\nu)} = \left(1 - \frac{\omega_p^2}{\omega^2 + \nu^2}\right) - j\frac{\nu\,\omega_p^2}{\omega(\omega^2 + \nu^2)} = \varepsilon_r - j\varepsilon_i \qquad (2.15)$$

where $\nu = 1/\tau$ is the collision frequency, τ is the scattering time between collisions, and ω_p is the plasma frequency that is expressed as:

$$\omega_p = \sqrt{\frac{Ne^2}{\varepsilon_0 m_e}} \qquad (2.16)$$

in which e and m_e are the electron charge and mass respectively and N is the concentration of free electrons. Plasma frequencies for metals are in the visible and

ultraviolet ranges. Typical values of plasma and collision frequencies are $\omega_p = 0.2300 \times 10^{12}\,\text{Hz}$ and $\nu = 5.510 \times 10^{12}\,\text{Hz}$ for gold and $\omega_p = 0.2100 \times 10^{16}\,\text{Hz}$ and $\nu = 4.450 \times 10^{12}\,\text{Hz}$ for silver, however impurities and other factors can affect these numbers so the literature does not have a unique set of values for these parameters. A comprehensive list of plasma and collision frequencies published by a variety of researchers for gold, silver and other noble metals can be found at [28].

Collisions are primarily between electrons and relatively large lattice vibrations (phonons), so for most metals at room temperature $\nu \ll \omega_p$ [29], which reduces (2.15) to approximately:

$$\varepsilon(\omega) \simeq 1 - \frac{\omega_p^2}{\omega^2} \tag{2.17}$$

This important expression shows that when $\omega < \omega_p$, then $\varepsilon(\omega) < 0$ which in part satisfies the conditions in (2.9). Therefore, surface plasmon polaritons can only exist below the plasma frequency. Equation (2.17) substituted into (2.6) gives a simplified Drude model for the plasmon wavenumber:

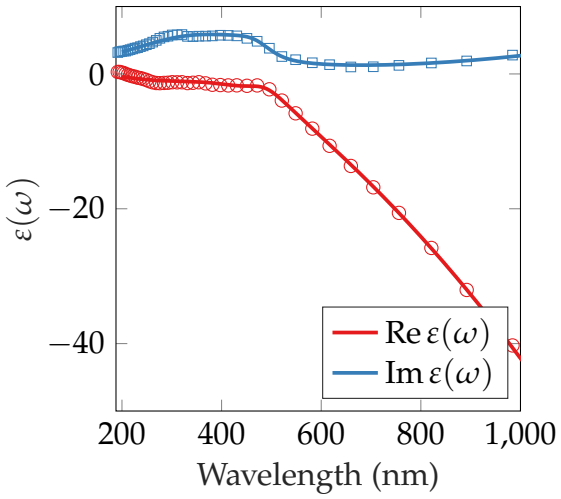
$$k_x \simeq k_1 \left(\frac{\omega^2 - \omega_p^2}{\omega^2 \varepsilon_{r1} + (\omega^2 - \omega_p^2)} \right)^{1/2} \tag{2.18}$$

The Drude model in the form given in (2.15) is widely used in nanoantenna analytical and numerical calculations for determining the permittivity, and therefore the propagation constant, of the plasmon polariton of noble metals. However, this free electron model fails at frequencies higher than about 850 THz [30], due to bound charge effects, and must be modified by incorporating additional terms [29]. This becomes a tedious process but reasonable accuracy can be achieved.

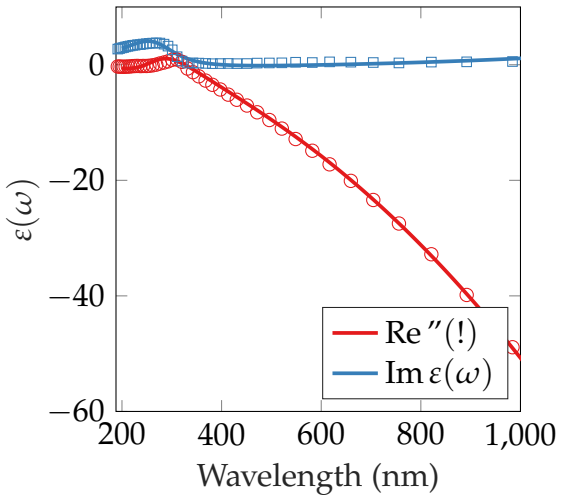
Another option, preferred by the authors, is to use measured data [31] with a partial fraction fit [32]. Fig. 2.5 shows the dielectric function for gold and silver over the frequency range from 0.5000 eV~120 THz to 6.500 eV~1.570 PHz obtained from measured data via a partial fraction fit. The solid lines are the partial fraction fit for the real part ε', and the dashed lines are for the imaginary part ε'', of the dielectric function.

The significance of the data in Fig. 2.5 becomes clearer in Fig. 2.6 where the surface plasmon dispersion curve for frequency versus the propagation k_x', attenuation k_x'', and constants at an air-silver boundary are plotted. The lower part of the curve is the non-radiative surface plasmon region followed above by an anomalous dispersion region and above that a Brewster mode radiative region. The slanted dashed line is the free-space light line $k_0 = \omega/c$. The Brewster mode portion of the dispersion curve is in the fast wave region to the left of the light line. It is the locus of the values of k_x of the plane wave incident on the boundary at the Brewster angle. These waves do not set up a surface wave, but rather carry their power directly into the metal without reflection. Below the Brewster mode section is a region of anomalous dispersion, often described in the literature as 'backbending', which has a group velocity that is virtually zero, and is accompanied by high loss.

The surface plasmon portion of the dispersion curve is of primary importance to nanoantenna design. The plasmon region lies to the right of the light line and therefore, it is a slow wave since the phase velocity here is less than the speed of light. Because surface plasmons have a shorter wavelength than light they are prevented from radiating from a flat planar surface. Of particular significance is the decreasing slope of the dispersion curve, tending toward horizontal near where it joins the anomalous dispersion segment. In the plasmon region the sur-



(a)



(b)

Figure 2.5: Dielectric function for (a) Gold and (b) Silver as a function of photon energy $E = hf$ where f is frequency and $h = 6.626 \times 10^{34}\,\mathrm{kg\,m^2/s}$ is Planck's constant. The solid lines are the partial fraction fit for ε' and the dashed line is for ε''. The circles and squares are from measured data originally listed in [1]

face plasmon propagation constant increases dramatically with a small increase in frequency. The horizontal dashed line at 380 nm intersects the light line at approximately $k_0 = 17\,\mathrm{rad/\mu m}$ while at the same wavelength $k'_x = 21\,\mathrm{rad/\mu m}$, which means the surface plasmon wavelength can be much less than the free-

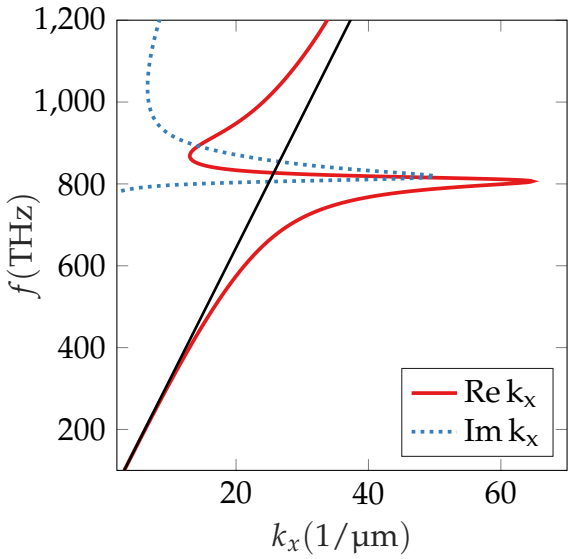


Figure 2.6: SPP dispersion curve for ω versus Re(k_x) and Im(k_x) for an glass-silver interface. The lower part of the curve is the non-radiative surface plasmon region which is capped by a high-loss backbending region, and a Brewster mode radiative region above

space wavelength. Therefore, a resonant half-wavelength nano-dipole antenna can be much smaller than a half-wavelength measured in air at the same frequency. As the plasmon wavelength decreases near the anomalous dispersion region, the propagation constant approaches a limiting value. The Drude model provides the estimate for this limiting frequency

$$\omega_{sp} \simeq \omega_p / \sqrt{1 + \varepsilon_{r1}}, \tag{2.19}$$

obtained by setting the denominator in (2.18) equal to zero, which is equivalent to assuming the propagation constant in (2.18) approaches infinity. This is not what happens to the propagation constant, as can be seen in Fig. 2.6 where real data is used, but it does provide a reasonable estimate of the limiting frequency. nanoantennas must be designed to operate below this frequency, thereby avoid-

ing a group velocity approaching zero and dramatically increased losses.

Although several key aspects of wave propagation on nano-structures have been presented here, all analysis has been performed assuming propagation is on a surface between two half spaces. However, most nanoantennas have air above and are etched in a noble metal, such as gold, on a dielectric substrate, such as silicon dioxide. The metal antenna, therefore lies between two different materials. nanoantennas are typically excited at a corner or edge where both upper and lower surfaces are exposed to the incoming electromagnetic field. Because the plasmon propagation constant (2.7) is a function of the permittivity in the regions above and below an interface, it will be helpful to understand how different phase velocities (2.10) and wavelengths (2.11) on the two sides of a nanoantenna will affect its performance. This issue is briefly addressed in the following paragraphs.

2.4 Two Boundary Structures

Surface plasmon polaritons on a metal film have properties that do not exist on an interface between metal-dielectric half spaces. To analyze the mode configuration on a metal strip of thickness t, consider the configuration shown in Fig. 2.7 where the dielectrics are assumed to extend from the surface of the metal to infinity and possess different permittivities, ε_1 and ε_2. The permittivity of the metal is ε_m and all three regions have the permeability of free-space. The waves in each region are assumed to be TM polarized with a magnetic field component in the y-direction in each region given by:

$$\mathbf{H}_1 = \hat{\mathbf{y}}\, H_{y1}\, \mathrm{e}^{-\mathrm{A}_{z1}z - jk_x x}, \qquad \text{dielectric region 1}(z \geq t) \tag{2.20a}$$

$$\mathbf{H}_m = \hat{\mathbf{y}} \left(H_{ya}\, \mathrm{e}^{+\mathrm{A}_{zm}z} + H_{yb}\, \mathrm{e}^{-\mathrm{A}_{zm}z} \right) \mathrm{e}^{-jk_x x}, \qquad \text{metal region m}(0 \leq z \leq t) \tag{2.20b}$$

$$\mathbf{H}_2 = \hat{\mathbf{y}}\, H_{y2}\, \mathrm{e}^{-\mathrm{A}_{z3}z - jk_x x}, \qquad \text{dielectric region 2}(z \leq 0) \tag{2.20c}$$

where the dispersion relations for the three regions are,

$$\mathrm{A}_{zi}^2 = -k_{zi}^2 = k_x^2 - k_0^2 \varepsilon_i, \qquad \text{for } i = 1, 2, m \tag{2.21}$$

and the boundary conditions satisfied by $\mathbf{H}_m$ at each interface are:

$$H_{ym} = H_{yi} \tag{2.22a}$$

$$\frac{1}{\varepsilon_m} \frac{\partial H_{ym}}{\partial z} = \frac{1}{\varepsilon_i} \frac{\partial H_{yi}}{\partial z} \tag{2.22b}$$

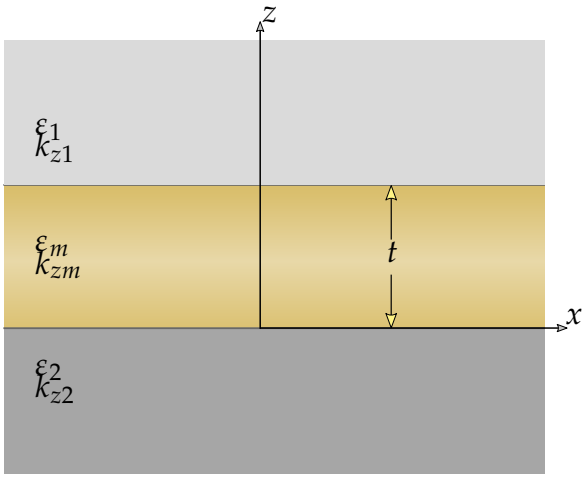


Figure 2.7: Thin metallic film between two dielectric regions with different permittivities.

where $i = 1, 2$. Each of these conditions are applied to (2.20) at each boundary yielding a set of four homogeneous linear equations which can be combined to produce:

$$\frac{H_{yb}}{H_{ya}} = \frac{\left(\frac{\varepsilon_2 \, \mathrm{A}_{zm}}{\varepsilon_m \, \mathrm{A}_{z2}} - 1\right)}{\left(\frac{\varepsilon_2 \, \mathrm{A}_{zm}}{\varepsilon_m \, \mathrm{A}_{z2}} + 1\right)} \tag{2.23a}$$

$$\frac{H_{yb} \, \mathrm{e}^{-2\mathrm{A}_{zm}t}}{H_{ya}} = \frac{\left(\frac{\varepsilon_1 \, \mathrm{A}_{zm}}{\varepsilon_m \, \mathrm{A}_{z1}} + 1\right)}{\left(\frac{\varepsilon_1 \, \mathrm{A}_{zm}}{\varepsilon_m \, \mathrm{A}_{z1}} - 1\right)} \tag{2.23b}$$

The equations in (2.23) are equated to produce the three dispersion relation:

$$\left(\frac{\varepsilon_m \, \mathrm{A}_{z1}}{\varepsilon_1 \, \mathrm{A}_{zm}} + 1\right)\left(\frac{\varepsilon_m \, \mathrm{A}_{z2}}{\varepsilon_2 \, \mathrm{A}_{zm}} + 1\right) = \left(\frac{\varepsilon_m \, \mathrm{A}_{z1}}{\varepsilon_1 \, \mathrm{A}_{zm}} - 1\right)\left(\frac{\varepsilon_m \, \mathrm{A}_{z2}}{\varepsilon_2 \, \mathrm{A}_{zm}} - 1\right) \mathrm{e}^{-2\mathrm{A}_{zm}t} \tag{2.24}$$

for the metal layer sandwiched between two different dielectrics. Equation (2.24) can be verified by allowing the metal thickness to become large, $\mathrm{A}_{zm}t \to \infty$, in which case the right hand side goes to zero. This yields a pair of equations:

$$\frac{\varepsilon_m \mathrm{A}_{z1}}{\varepsilon_1 \, \mathrm{A}_{zm}} + 1 = 0 \tag{2.25a}$$

$$\frac{\varepsilon_m \mathrm{A}_{z2}}{\varepsilon_2 \, \mathrm{A}_{zm}} + 1 = 0, \tag{2.25b}$$

recognized from (2.4) to be the dispersion relations for surface plasmon polaritons at the boundaries between two individual dielectric-metal half-spaces.

If ε_1 and ε_2 are positive and $\varepsilon_m < 0$, the right hand side of (2.24) is positive, which means the left hand side must also be positive. This is possible if, $\left(\frac{\varepsilon_m \, \mathrm{A}_{z1}}{\varepsilon_1 \, \mathrm{A}_{zm}} + 1\right)$ and $\left(\frac{\varepsilon_m \, \mathrm{A}_{z2}}{\varepsilon_2 \, \mathrm{A}_{zm}} + 1\right)$ are both positive or both negative. Under the assumption that both are negative, that $\varepsilon_1 > \varepsilon_2$, and $|\varepsilon_m| > \varepsilon_1$ the characteristic

equation (2.24) becomes [33],

$$A_{zm}t = \coth^{-1}\left(\frac{|\varepsilon_m|\, A_{z1}}{\varepsilon_1\, A_{zm}}\right) + \coth^{-1}\left(\frac{|\varepsilon_m|\, A_{z2}}{\varepsilon_2\, A_{zm}}\right). \tag{2.26}$$

Similarly, assuming that both expressions in brackets on the left hand side of (2.24) are positive and $|\varepsilon_m| < \frac{\varepsilon_1\varepsilon_2}{\varepsilon_1-\varepsilon_2}$ the characteristic equation becomes [33],

$$A_{zm}t = \tanh^{-1}\left(\frac{|\varepsilon_m|\, A_{z1}}{\varepsilon_1\, A_{zm}}\right) + \tanh^{-1}\left(\frac{|\varepsilon_m|\, A_{z2}}{\varepsilon_2\, A_{zm}}\right). \tag{2.27}$$

Equations (2.21) together with (2.26) or (2.27) can be used to solve for the wavenumbers in the even and odd mode cases, referring respectively to the symmetric or asymmetric form of the tangential (z-directed) component of the electric field about the center of the metal layer [34].

When the metal layer is thick, the waves on the two sides of the metal slab do not interact, as indicated in (2.25). However, as the slab becomes thinner the electromagnetic fields on the two sides do interact and the frequency splits into a low frequency even mode (2.26) and a higher frequency odd mode (2.27). Although the tangential electric field is not entirely symmetric in the metal region when $\varepsilon_1 \neq \varepsilon_2$, more of the field lies inside the metal in the even mode case (2.26) than it does in the basically asymmetric odd mode case (2.27). As the metal layer becomes thinner, the damping of the odd mode decreases, approximately as the square of the thickness, allowing this mode to travel a greater distance than a plasmon polariton on the interface between two half spaces. However, the basically symmetric mode damping increases with decreasing thickness, allowing propagation distances one or two orders of magnitude greater than those of thick metal construction [35]. The physical reason attributed to this phenomena is that

in the asymmetric case, where the null is in the center of the metal, more of the field resides outside as the thickness of the metal decreases therefore, lessening the presence of the tangential electric field in the metal thereby reducing Joule heat which in turn decreases the electron collision rate resulting in less damping of the wave. This analysis suggests that the efficiency of an optical antenna can be improved by decreasing the thickness of the metallic material with which it is constructed and operating in the higher odd mode frequency range.

When the metal width is finite and the environment is symmetric ($\varepsilon_1 = \varepsilon_2$), there are four fundamental modes and a number of higher order modes. These modes can be classified according to the cross-sectional width and thickness, similar to a rectangular waveguide and will become successively cut off as the thickness decreases leaving one Gaussian like mode. Efficient optical fiber coupling to this plasmonic mode in a metal guide has been demonstrated [36], raising the possibility of similar coupling between optical antennas and metal strip plasmonic waveguides designed with the appropriate width and thickness so as to operate in this particular mode. Closer to the nanoantenna case, in an asymmetric environment ($\varepsilon_1 \neq \varepsilon_2$) there is no purely TM mode. It has been shown that modes in this case can change their symmetry properties with changes in both transverse directions, width and thickness, and that all modes show a cutoff height that increases both with decreasing width and increasing dielectric mismatch between the sub- and superstrate [36, 37]. Unfortunately, this includes air-metal-glass configurations commonly used in the design of optical antennas and waveguides. Thus the complicated spatial profiles shown in modes in an asymmetric environment prohibit efficient excitation techniques [24, 38].

In this section it has been shown that in the optical frequency range electromagnetic wave propagation, which is via a surface plasmon polariton, can take

place on an air-dielectric boundary if one of the two materials has a dielectric constant with a negative real part obeying the conditions in (2.9). The dispersion equation for this wave is given in (2.6) and presented graphically in Fig. 2.6, which shows that a surface plasmon is a slow wave, i.e. slower than the speed of light in the dielectric medium. The Drude model, which is relatively accurate below, and not too close to, the plasma frequency ω_p and useful for analytical calculations, was given in (2.15). This model, or its approximation in (2.17), is commonly used to obtain approximate velocity (2.10), wavelength (2.11), propagation distance (2.12), and exponential decay distance away from the boundary (2.13) and (2.14) of the plasmon wave. It also provides an estimate of maximum frequency (2.19) at which a plasmon wave can propagate. Although Fig. 2.6 shows that the surface plasmon wavelength can be much less than the free-space wavelength, close to the maximum frequency the plasmon wave will experience severe attenuation.

Because optical antennas are operated in a frequency range which is at the overlapping boundary between electromagnetic and quantum theories, in the paragraphs above, some of the quantum justification for the unusual behavior of materials in the near infrared and optical domains has been presented. This discussion was then followed by an asymmetric two boundary case development of the dispersion equations modeling a typical optical antenna structure in cross-section, showing that the interaction between the plasmon waves on the two boundaries breaks into symmetric and asymmetric modes, (2.26) and (2.27). The symmetric mode is preferred when coupling to another device and the asymmetric mode is preferred for long distance propagation in a plasmon waveguide. This section then concluded with a discussion of the finite width metal guide that has been shown to have a complicated mode structure, hampering efforts

to obtain a mathematical expression for impedance matching. Current attempts at obtaining a mathematical model and along with design methods that improve impedance matching and enhance optical antenna radiation will be covered in the individual antenna discussions that follow.

2.5 Design of Optical Antennas

The plasmonic nature of nanoantenna operation prevents the designer from describing antenna performance with formulas that have been developed over many decades at microwave and lower frequencies. Dielectric properties of metals are only one factor preventing a direct correspondence between established antenna theory and actual optical antenna operation. A traditional antenna concept that the wave vector for the current is the same as the free-space wave vector must be abandoned, dispersion effects must be accounted for and losses not generally expected in a microwave frequency 'perfect conductor' become dependent on the size and shape of the antenna. The optical antenna, although not a traditional perfect conductor, cannot be analyzed as though it is a pure dielectric rod as would be the case for an optical fiber. But even if an optical antenna could be modeled as a pure dielectric, the tips of a rod represent the famous unsolved dielectric wedge problem. As described above, the fast wave current on a perfect conductor must be replaced by a slow wave plasmon current tightly coupled to the antenna surface. Radiation on an optical dipole still occurs at the tips of the rods and can still be viewed as current induced. The primary difficulty that must be faced when attempting an analytical analysis is that for practical reasons the design requires placing the antenna on an interface between a dielectric below, such as a silicon dioxide substrate, and air above. Although numerical techniques can be used to accurately analyze the characteristics of the structure, similar to

microstrip patch antennas, no exact analytical methods are available. Analytical approximations and numerical attempts to obtain the basic characteristics of an optical antenna, the resonant frequency, input impedance, gain, directivity and efficiency, are presented in this section.

2.5.1 Dipole and Patch Antennas

Size and shape can affect the design frequency of an optical antenna in ways not typically expected in microwave frequency antenna engineering. For example a rectangular microstrip patch is expected to resonate when the incident field is polarized in the direction of the $\lambda/2$ dimension of the patch, while the transverse dimension is flexibly left to something greater or less than $\lambda/2$. However, as shown in Fig. 2.9 both the width and length of a metal plate must be considered in designing an optical antenna. A narrow nanodipole resonating at a particular frequency, as shown in Fig. 2.8, does not maintain resonance when widened. However, the first resonance can be regained by changing both the length and width of the dipole as shown in Fig. 2.9c. The mechanism causing this phenomena has to do with the manner in which surface plasmons are excited.

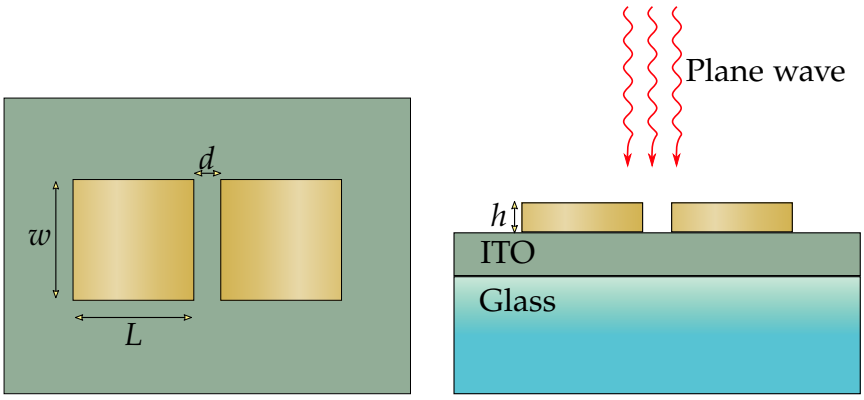
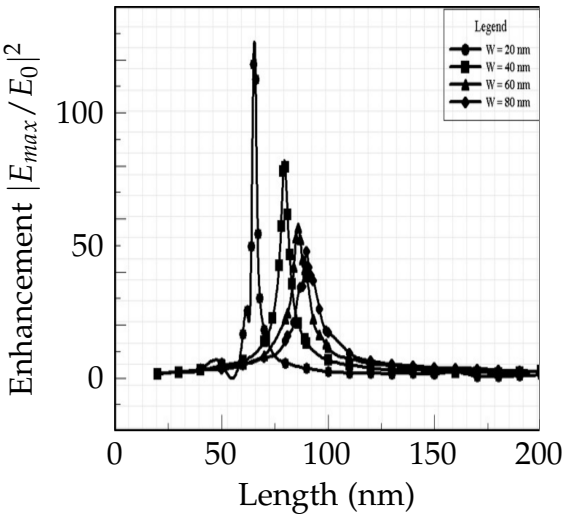
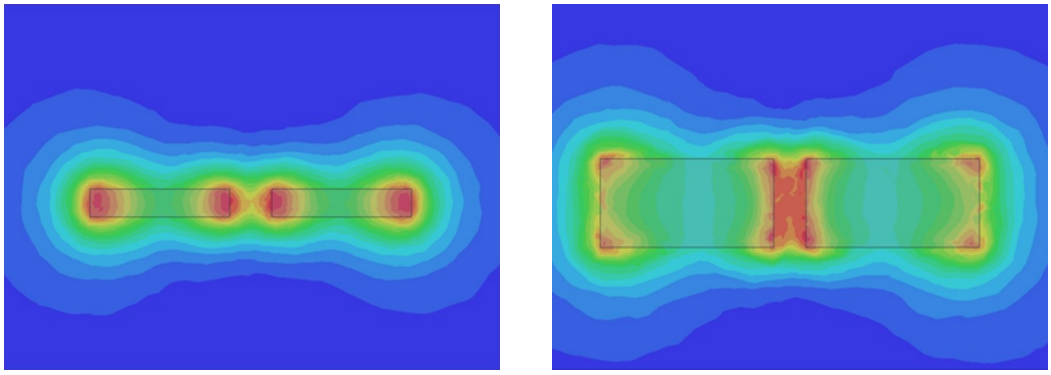


Figure 2.8: A gold dipole residing over a glass substrate.



(a)



(b) (c)

Figure 2.9: (a) Maximum enhancement of a gold nanodipole excited at 550 nm (b) narrow and (c) wide profile. A narrow structure does not maintain resonance when widened without also changing its length.

The incident wave with a free-space wavenumber transfers energy into the plasmon wave having a much larger wavenumber at locations where the interaction between the incident wave and the antenna produce the largest concentration of higher order modes, which will couple energy into the plasmon mode. This interaction is primarily seen on the corners of the dipole in Fig. 2.9c. Similar to plane waves in a waveguide, plasmon waves fan out from the corners taking

a circuitous route bouncing from the edges of the plate before phase matching produces a resonance condition. Further complicating the picture is the chemical potential and the 'lighting rod' effect whereby opposite charges form a high concentration of the field across the antenna center gap. However, plasmons in the narrower dipole, Fig. 2.9b, follow a shorter, more direct route to resonance than does the plasmon wave on the wider dipole Fig. 2.9c thereby experiencing less attenuation and ultimately creating a region of higher field enhancement in the gap than does the wide dipole. Fig. 2.9a shows that as the dipole widens the resonant frequency shifts and the enhancement decreases. Enhancement is defined as the ratio of the total to the incident electric field in the gap region, although in some publications this ratio is squared and therefore, described as intensity enhancement.

The behavior of the narrower dipole shown in Fig. 2.9b is more closely aligned with what one would expect at microwave frequencies whereby increasing the length of each nanorod in plasmon half-wavelength intervals will produce successive resonance and antiresonance in the gap input impedance. Notice that the total length of the dipole is a full plasmon wavelength λ_p plus the gap separation distance, as opposed to the typical $\lambda_0/2$ dipole first resonance of a perfectly conducting microwave antenna with a negligible gap width. In terms of plasmon wavelengths, a first resonance does occur in an isolated $\lambda_p/2$ nanorod or a vertical $\lambda_p/4$ nanorod on a metal substrate [39].

Although an exact determination of the plasmonic wavelength in a dielectric nanorod is not mathematically feasible, an effective wavelength λ_{eff} based on the assumption that the antenna is a thin dielectric cylinder operating in the TM_0 cylindrical waveguide mode has been derived in [40]. The antenna is immersed in a material with dielectric constant ε_s, and a Drude dielectric function is used

to describe the metal rod material. The apparent increase in the antenna length at the rod ends is accounted for by an additional reactance term. This scaling law is:

$$\lambda_{eff} = n_1 + n_2 \left(\frac{\lambda}{\lambda_p} \right) \tag{2.28}$$

where λ is the wavelength of the external region, λ_p is the metal plasma wavelength, and n_1, n_2 coefficients with dimensions of length that depend on antenna geometry and static dielectric properties, are

$$n_1 = 2\,\pi\,R \left[13.74 - 0.12 \frac{\varepsilon_\infty + 141.04\,\varepsilon_s}{\varepsilon_s} - 2\pi \right] \tag{2.29a}$$

$$n_2 = 0.24\,\pi\,R \frac{\sqrt{\varepsilon_\infty + 141.04\,\varepsilon_s}}{\varepsilon_s}. \tag{2.29b}$$

Here R is the radius of the cylinder, and ε_∞ is the dielectric constant correction in the Drude formula for $\omega \gg \omega_p$. For gold $\varepsilon_\infty \simeq 11$, $\lambda_p = 138\,\text{nm}$ and for silver $\varepsilon_\infty \simeq 3.5$, $\lambda_p = 135\,\text{nm}$ [40].

In the case of an optical patch or dipole antenna it is necessary to account for a phase shift acquired by the surface plasmon polariton upon reflection at rod, strip, or gap terminations. The resonance wavelength λ of metal strip resonators of thickness t and width w are determined approximately by [41]:

$$w \frac{2\pi}{\lambda} n_{slow} = m\pi - \phi \tag{2.30}$$

where n_{slow} is the real part of the mode index of surface plasmon polaritons propagating in a metal film with the same thickness t as the strip, $m = 1, 2, 3, \ldots$ is the order of the resonance, and ϕ is the phase of the reflection coefficient at strip terminations. For a symmetric structure where $\varepsilon_d = \varepsilon_1 = \varepsilon_2$, once t and ε_m

are selected and the phase shift ϕ is determined, n_{slow} can be obtained from (2.21) and (2.25) or for thin strips, one can use the approximation [42]:

$$n_{slow} \simeq \sqrt{n_d^2 + \frac{4\, n_d^4}{k_0^2\, t^2\, n_m^4}} \tag{2.31}$$

where n_d and n_m are the indices of refraction of the dielectric and metal and $k_0 = 2\pi/\lambda$. The wavelength of the slow surface plasmon polariton is then given by $\lambda_{slow} = \lambda/n_{slow}$.

Analytical models are preferable because clear trends in antenna performance can be seen by adjusting mathematical parameters that represent physical quantities such as the antenna width or length and the excitation frequency. On the other hand, numerical techniques can provide both the accuracy and flexibility needed to model complex antenna structures. However, purely mathematical solutions can become clouded with long complex expressions and special functions while a purely numerical analysis provided by the powerful commercial codes that have become available these last few years, while furnishing useful visual data in terms of graphs and charts, do not always provide a way to understand the multifaceted interaction between the variables of an antenna system. Recently a compromise approach for optical antennas has been advanced, advocated primarily by Engheta and Alù [43, 44, 45, 46], whereby the antenna and field excitation are replaced by an equivalent lumped circuit model based on numerical analysis data. The basic model, shown in Fig. 2.10, is centered on a Thévenin equivalent circuit for the antenna with a capacitance accounting for a non-negligible gap impedance Z_g along with the intrinsic impedance Z_a of the dipole. The parallel combination of these two impedances forms the input impedance of the circuit. The input impedance is calculated, by driving the

antenna at the gap with an arbitrary source voltage V_g and calculating, with full-wave simulations, the displacement current I_g flowing across the arms at the terminals in the region of the gap $Z_{in} = V_g / I_g$.

The intrinsic antenna impedance Z_a is extracted from the input impedance by first taking the parallel combination of the gap and load impedances according to:

$$Z_{in} = \frac{1}{1/Z_a + 1/Z_g}, \quad Z_g = \frac{1}{j\omega C}, \quad \text{where } C = \varepsilon_0 S/g \tag{2.32}$$

where S and g are respectively the gap cross-sectional area and gap height. Setting $Z_a = R_a + jX_a$ and $Z_{in} = R_0 + jX_0$, (2.32) is rearranged to give:

$$R_a = \frac{R_0}{1 + \omega C \left(2X_0 + \omega C \left(R_0^2 + X_0^2\right)\right)} \tag{2.33a}$$

$$X_a = \frac{X_0 + \omega C \left(R_0^2 + X_0^2\right)}{1 + \omega C \left(2X_0 + \omega C \left(R_0^2 + X_0^2\right)\right)}. \tag{2.33b}$$

Since full-wave simulations have provided the Z_{in} components R_0 and X_0, the dipole intrinsic impedance is determined mathematically from (2.33). The intrinsic impedance Z_a, which is unaffected by antenna loading either by a transmission line or by altering the gap capacitance, is now completely determined.

The antenna is resonant when the input reactance $X_0 = 0$ which from (2.33b) is:

$$\omega_0 = \frac{X_a}{C(R_a^2 + X_a^2)} \tag{2.34}$$

The 'open circuit' or first resonance has been shown to have a very low value, whereas the 'short circuit' or antiresonance has a high impedance, on the order of kilo-ohms. The feed line or other loading conditions will determine which operating frequency is selected. In the event of loading, the capacitance becomes

$C = \varepsilon_L S/g$, the new resistance and reactance terms in the input impedance $Z_{in} = R_{in} + jX_{in}$ are again found numerically and the equivalent dielectric constant ε_L can be calculated from (2.33b). Extra circuit elements can be added to account for the transmission feed line or a quantum radiator such as is the case in molecular spectroscopy, or a different circuit can be created for single quantum radiators without a gap [46]. Efficiency is an important measure used to characterize antenna performance. The overall antenna efficiency can be broken down into the product of several measures of efficiency, the most important of which are the radiation efficiency and the reflection efficiency, $\eta = \eta_{rad}\eta_{ref}$, where the maximum value of each efficiency is one. The reflection efficiency determines how well the transmission feed line characteristic impedance is matched to the antenna input impedance. Its value, given in terms of the transmission line reflection coefficient Γ, is $\eta_{ref} = 1 - |\Gamma|^2$ where $|\Gamma| \leq 1$. The radiation efficiency is ratio of the total power radiated P_{rad} to the power received P_{in} by the antenna. The input power is assumed to be the sum of the radiated power and the power losses on the antenna. Provided the antenna is fed at the antenna maximum current point, which is most often the case with optical antennas operating at

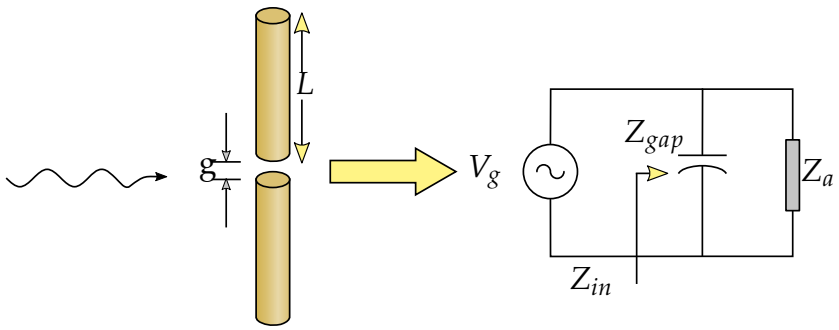


Figure 2.10: Optical antenna with equivalent circuit model. The driving voltage could be an incident electromagnetic field, a transmission line connected to the antenna, or a quantum emitter.

or below the lowest resonant frequency. At the first antiresonance, the radiation efficiency can be expressed in terms the radiation resistance R_{rad} and the ohmic loss resistance R_L experienced by the current flow in the antenna elements. The radiation efficiency is therefore:

$$\eta_{rad} = \frac{R_{rad}}{R_{rad} + R_{in}\sin^2(\pi\, L_{eff}/\lambda_{eff})}; \quad R_{rad} = 2P_{rad}/I_{max}^2. \tag{2.35}$$

The ohmic loss resistance R_{in} is the real part of the input impedance. The optical dipole radiation resistance can be determined analytically by first hypothesizing the current at resonance. Taking into account the effective length and wavelength of an optical antenna, L_{eff} and λ_{eff} respectively which can be calculated using (2.28), the standard microwave expression for a dipole current is modified to become [44],

$$I(z) = \frac{I_0 \sin\left[\pi \dfrac{L_{eff} - 2|z|}{\lambda_{eff}}\right]}{\sin\left(\dfrac{\pi\, L_{eff}}{\lambda_{eff}}\right)} \tag{2.36}$$

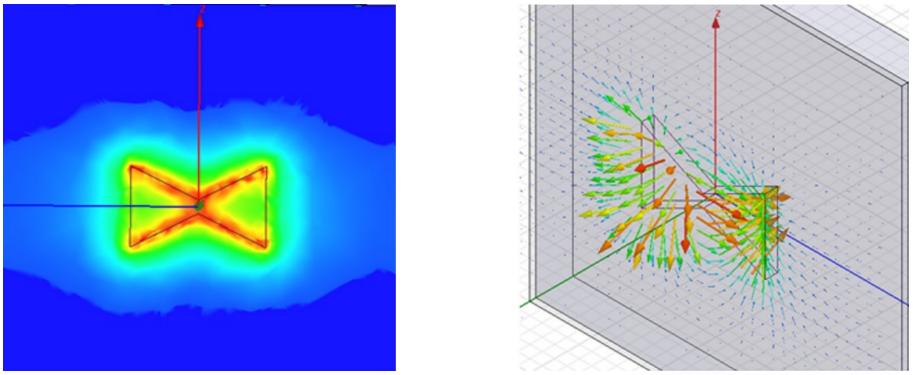
where I_0 is the numerical displacement current at the feed point $z = 0$. This current (2.36) is used to find an analytical expression for the power radiated by the optical dipole P_{rad} which is then substituted into (2.35) to find the efficiency.

Antenna directivity defined as D_0 = (time average power density maximized with angle)/(average total power radiated through a sphere), can again be analytically calculated with the formulas above as can gain which is the product of the efficiency times the directivity, $G_0 = \eta_{rad} D_0$. Keep in mind that this exercise has been a combination of numerical and analytical calculations, forming a hybrid approach that seems best suited for understanding complex interactions as well as for overall engineering optical antenna design. Although the example of

a dipole has been used here, many other structures have been successfully investigated using this equivalent technique [45, 46]. For design data for the dipole based on measurements the reader is referred to references [47, 48, 49].

2.5.2 Bowtie Antenna

Compared to a dipole antenna, the bowtie antenna has the advantage of broadband operation yet it is a simple design. The feed point of a bowtie antenna is a gap at the vertex of two tip to tip apposal triangular electrically conducting arms. If extended to infinity while the angles of the triangle vertices remain unchanged, this antenna would be frequency independent. When the incident field polarization is perpendicular to the gap, the antenna can be driven into resonance. Fig. 2.11 shows the intensity distribution for a 520 nm resonant bowtie design made out of gold on a SiO_2 substrate and excited by a normally incident plane wave. The numerical simulation in Fig. 2.11a shows that a high field intensity distribution occurs not only at corners and in the gap region but also along each of the sides, likely due to the sharp edge design. In Fig. 2.11b a plot of the electric field vectors immediately above the surface of antenna indicate the largest field components are just above the gap region but tend to decrease in amplitude as they spread out and terminate at the wider ends of the antenna arms where a much lower intensity is seen in Fig. 2.11a. The intensity of the field in the bowtie gap has been shown to be strongly dependent upon the gap width [47]. For the design presented in [47], starting on the order of 50 nm the resonant intensity climbs exponentially as the gap separation narrows, however as the gap separation widens the intensity tends to level out at around 100 nm. Although very little work has been carried out concerning mathematical design criteria for the construction of on-surface optical antennas, optimal design dimensions can



(a) (b)

Figure 2.11: Bowtie antenna with (a) highest field intensity shown in red around the edges and corners of the triangular arms and (b) electric field vectors where the largest amplitude field is displayed in red occurs in the gap between the two arms.

be obtained through published data gathered in experiments. For optical frequency design data based on measurements for the bowtie antenna the reader is referred to reference [48]. Studies have also been carried out to determine optimum bow angle, the outside angle formed by the edges of the two apposal arms of a bowtie antenna. It has been reported that the strongest enhancement can be obtained with a bow angle of 90° [48]. Unlike microstrip antennas printed on a thin substrate with a metal backing focusing radiation into the air region above, optical antennas are a single metal radiator lying on a dielectric-air interface. There can be no ground plane below because there are no perfect conductors at optical frequencies. Although radiation is in both directions, most of the energy radiated from an optical antenna tends to couple into the higher density substrate. In the microwave case a TM_0 mode surface wave can be launched into the substrate because this mode has no cutoff wavelength, but the fundamental mode in a typical microwave patch antenna is not conducive to launching TM

modes. As long as the substrate is thin, a microwave patch antenna suffers little substrate induced loss. As the substrate thickens, energy from the microwave patch antenna is coupled into surface wave modes, which seriously degrades the efficiency of an antenna. However, an optical antenna radiates in both directions regardless of the substrate thickness, making it in any case less efficient than its microwave counterpart. One measurement study carried out in the infrared range [48] concluded that the dipole and bowtie efficiencies are respectively 20 % and 30 % whereas a typical microstrip patch antenna has an efficiency of 60 %. A similar conclusion was reached in a more recent study [46] that included the optical dipole, bowtie, and dimer constructions. A dimer, which has two pancake shaped arms, has an efficiency level close to the bowtie design. Numerical investigations of the optical Yagi-Uda design show a strong emission into the substrate and a much smaller emission lobe into the air side of the surface [50] [51]. Although most of the electromagnetic radiation is emitted into the critical angle below, the above surface emission still shows a strong directivity.

2.5.3 Yagi-Uda Antenna

The dipole and bowtie nanoantennas described above are most often used in applications where a wide beamwidth or broadband transmission is the primary objective. However as nanotechnology progresses, achieving highly directional beams in the optical regime will be necessary in order to, for instance, transmit signals from nanocircuits to a receiver above the surface with a minimum power load on the electronic system. Another significant application is the possibility of enhancing and directing the radiation from a nano-emitter, which will be discussed in the *Applications* section below. At the microwave level the Yagi-Uda antenna shown in Fig. 2.12 has proven to be highly directional for radiation of ra-

dio waves yet simplistic in construction. One would expect that basically similar design principles can be used to obtain directionality with an optical Yagi-Uda design. Over the last few years, this has been shown to be the case primarily due to the fact that the elements of the Yagi antenna are nanorods, the properties of which have been the subject of extensive study by the optics community over the last thirty years. The essential Yagi-Uda design consists of a resonant

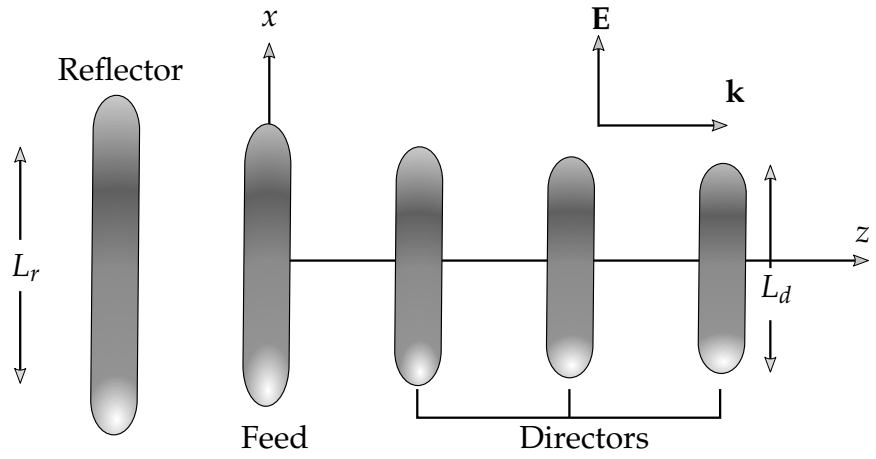


Figure 2.12: Yagi-Uda optical antenna constructed with nanorods.

driven element flanked on one side by a longer passive reflector element and on the other by a series of shorter director elements, all of which lie in the same plane. Researchers have studied a variety of element types for the Yagi, including spheres and prolate spheroids, but low efficiency and difficulties in achieving smooth surfaces and precise positioning on the nano level has focused attention on the classical cylindrical rod design. In addition to surface smoothness and the precise rod length and positioning needed for optimum array performance, other difficulties encountered in the optical regime are the loss of energy coupled into the substrate, the difference in wavelength between free-space and the metallic

rods, and field interaction with the air-substrate boundary causing a phase difference between surface plasmons on the two sides of the antenna. Naturally these effects can be modeled by pure numerical calculations but preventing unwanted consequences such as the substrate energy loss are still left open for future research. However detailed design rules have been developed [50] for the optical Yagi. As a rough estimate high directivity can be obtained with separation distances of about $0.25\,\lambda$ between the feed and the reflector and about $0.3\,\lambda$ between the feed and the director and between each of the director elements [51]. The length of the feed element is the effective resonant wavelength λ_{eff} of the antenna rods which is related to the incident free-space wavelength λ by a simple relation given in (2.28) above with an accompanying discussion.

A scanning near-field optical microscope (SNOM) has been used to simultaneously image the amplitude and phase of the normal electric field component on a working Yagi-Uda nanoantenna [52]. This data was processed to produce a time evolution animation of the fields of the antenna during reception, providing an unusual opportunity to check time domain numerical simulations of the received fields. These measurements showed that illumination of the antenna from the forward direction resulted in constructive interference of scattered light by the antenna elements which leads to a strong field enhancement in the feed element. Measurement data taken when illumination is from the rear of the antenna reveal that the destructive interference by the feed element suppresses strong fields.

2.5.4 Log-Periodic Antenna

To date, most theoretical and experimental studies on directional optical antennas have focused on the Yagi-Uda design, shown in Fig. 2.12, which pro-

vides good directivity for a specific design frequency with a limited bandwidth. A possibility of improving both bandwidth and directivity may be met with a log-periodic antenna design. The name is derived from a progressive spacing between the elements of the antenna based on a logarithmic scale. A number of log-periodic design schemes have been employed at microwave frequencies. At optical frequencies several graded size and spacing designs have been proposed with elements varying in shape from spheres and ellipses to the classical elongated rod. Due to difficulties in creating precise size and shapes at the nano level most implementations are only approximations of the true log-periodic requirement.

In a log-periodic antenna arrangement [53], the arm lengths l_n, spacing between arms R_n and rod diameters d_n, increase logarithmically as defined by the inverse of the geometric ratio τ as given by:

$$\frac{1}{\tau} = \frac{l_{n+1}}{l_n} = \frac{R_{n+1}}{R_n} = \frac{d_n}{d_{n+1}}. \tag{2.37}$$

Another parameter associated with the array, but independent of the geometric ratio, is the scaling factor. The scaling factor is often expressed normalized by the element length and as such it is given by $\sigma = d_n/2l_n$. The half angle at the antenna apex is determined by the parameters τ and σ according to $\mathrm{A} = \tan^{-1}\left[(1-\tau)/4\sigma\right]$.

In one study [54], measurements were carried out on an optical log-periodic antenna with gold elements on a glass substrate strictly following the design criteria above. However unlike other optical array designs, a gap was placed in the middle of each arm in order to enhance the field produced by plasmonic action. It was observed that the directionality of the antenna reaches a maximum with

$N = 10$ elements but the optimum field enhancement value of $|E_{FB}|^2$ for the forward beam electric field is obtained between $N = 6$ and $N = 10$. A decrease of the field enhancement values, defined as $|E_{total}|^2/|E_{inc}|^2$ with increasing numbers of elements is associated with increasing plasmonic losses incurred in the addition of elements. Adding a gap did not affect the antenna patterns, however a significant increase in beam intensity was observed. As with other on-surface plasmonic antennas, a significant portion of the radiated power is coupled into the substrate.

One of the drawbacks of the log-periodic design is that the currents in the optical version have the same phase in each element. In addition, the elements are closely spaced producing a phase progression of the current in the direction of the longer elements. This produces an end-fire beam in the direction of the longer elements and interference effects in the pattern. The standard design for a log-periodic antenna calls for crisscrossing the feed between adjacent elements thereby adding a 180° phase shift to the terminal of each element. Since the shorter elements are spaced closer together and opposite in phase, very little energy is radiated from them and interference effects are minimal. The radiation pattern in this crisscross case tends to be toward the shorter elements [53]. In the previous paragraph, the design calling for dipoles rather than solid rods appears to be a step in the right direction. The plasmonics of optical frequency materials and construction methods on the nano level will determine the efficacy of adapting the standard microwave design where alternate dipoles are connected.

2.5.5 Purcell Factor

Light emission from an optical quantum emitter such as a quantum dot or a nano-crystal can be enhanced when it is in close proximity to an intense elec-

tromagnetic source. This effect was first discovered at radio frequency by E. M. Purcell, when he observed an increase in the probability of spontaneous emission of an emitter placed in a resonator [55]. However, controlling the degree of enhancement of a quantum emitter remains a major challenge. Optical microcavities or microresonators such as optical nanoantennas, generally used for emission enhancement can be mathematically quantified by what has become known as the Purcell factor F [56], defined as:

$$F = \frac{3}{4\pi^2}\frac{\lambda}{n} \times \frac{Q}{V} \tag{2.38}$$

where λ is the wavelength of the emitted photon, Q and V are the quality factor and mode volume respectively, and n is the refractive index of the surrounding medium. The Q corresponds to the temporal confinement while the mode volume V represents the spatial confinement. Although cavity based structures can provide very high $Q(\sim 10^4)$ [57], they are consequently narrow bandwidth devices. This spectral limitation is a problem for contemporary solid-state quantum emitters which, despite many efforts in research, still have a broad spectral emission due to stability issues [58].

Optical nanoantennas, on the other hand have a much lower $Q(\sim 100)$ [59] and hence are suitable for applications involving broadband sources. However, due to surface plasmon polaritons, they have the capability to highly confine the light source volume well beyond the diffraction limit [60, 61]. Coupling optical nanoantennas with quantum structures like quantum dots can result in highly directive quantum emitter intensity pattern, but with a mode volume as low as $.002\,(\lambda/n)^3$ [59] which corresponds to a very high Purcell factor of $\sim 5 \times 10^4$.

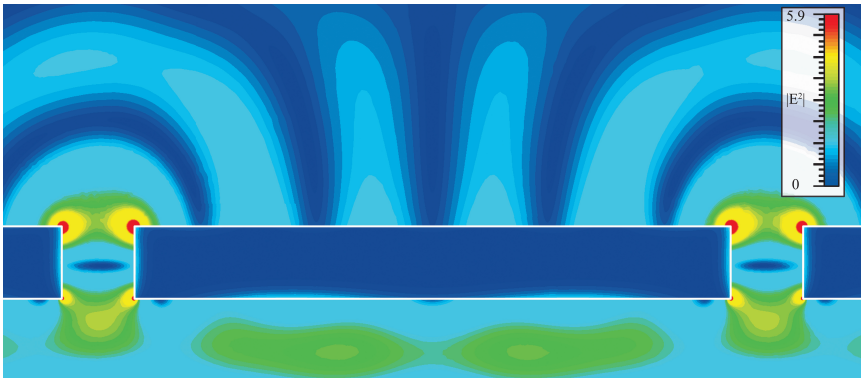
2.5.6 Aperture Antennas

So far in this section on-surface optical antennas have been emphasized primarily because there is considerable current scientific and industrial interest in the design of antennas for optical and near infrared communications. However, there is also significant industrial interest in antennas that are effective in focusing light to sub-wavelength dimensions. A primary application lies in the possibility of extending current photolithographic techniques to produce circuit chips on a sub-wavelength scale using relatively inexpensive optical equipment as compared to ion-beam etching machines. Nano-aperture antennas have been the key component in sub-wavelength lithographic research and also been a key element in the development of several other technologies that will be discussed in the *Applications* section. Below a brief treatment of the theory of optical aperture antennas is presented followed by an overview of selected designs that have received popular attention in the literature.

2.5.7 Optical Aperture Antenna Theory

Although important nano-level aperture optics was carried out in the study of fluorescence molecules in the 1980's [62], the report by [63] of dramatically enhanced transmission through sub-wavelength holes in metal films, was an event that spurred widespread interest in optical aperture antennas. Because this phenomena, now known as extraordinary optical transmission (EOT), seemed to fly in the face of the widely accepted classical electromagnetic analysis by [64] and [65], there has been considerable scientific interest in the subject reaching to the core of surface plasmon theory. According to the classical theory, the power transmitted through an aperture in an infinitely thin metallic screen scales as the inverse fourth power of the aperture size in terms of wavelengths. If the screen

is made to be of finite thickness, then an additional reduction in the transmitted field strength will result, due the below-cutoff dimensions of the slot. However, it was discovered, first experimentally and then confirmed by numerical analysis, as shown in Fig. 2.13, that for metals at optical frequencies there are situations where remarkable



(a)

Figure 2.13: Simulation showing a surface plasmon standing wave intensity ($|E|^2$) pattern on a thin silver plate with dielectric constant $\varepsilon = -18.242 + j1.195$ between two slots illuminated by a 660 nm plane wave from below. The center plate has a length $5/2\lambda_{spp}$ (1650 nm) and each aperture has a width and thickness of 200 nm.

amounts of energy pass through sub-wavelength slits and holes. Fig. 2.13 illustrates a cross-sectional view of a metallic sheet made of silver with two slots passing through the sheet. This nanostructure is excited by a plane wave from below. The magnitude of the field intensity shows a complicated structure with significant field at the slot corners, an interaction resulting in strong standing waves in the horizontal direction below, some weak field transmission through the slots, and a standing wave pattern on top of the center plate. Although [63] explained EOT in terms of surface plasmon polaritons, others [66, 67] questioned

the existence of surface plasmons and asserted that the observed phenomena could be explained by the existence of evanescent waves arising in a classical electromagnetics analysis not used by Bethe and Bouwkamp. However, a more detailed study that evolved through a series of technical papers confirmed the original EOT hypothesis [68, 17]. In the following, a purely mathematical analysis is presented revealing the nature of transmission through sub-wavelength holes including the surface plasmon, a space wave in the vicinity of the aperture, and a lateral wave behavior on the metal surface at distances over 100 surface plasmon wavelengths removed from the aperture. The importance of this analysis lies in the fact that surface plasmon radiation plays a more significant role in optical aperture radiation than does diffraction. Also because many applications require optical apertures with narrowly focused beams, it is important to know that significant energy is carried in surface plasmon polaritons on the surface of the metal on the side opposite the aperture excitation.

A thin slot aperture is modeled by placing a 2-D magnetic line current on the air-metal interface along the y-axis, into to the page, as shown in Fig. 2.2. This arrangement gives rise to a transverse magnetic (TM) field, which has three non-zero electromagnetic field components H_y, E_x and E_z. A straightforward analysis yields the magnetic field along the air-metal boundary:

$$H_y(x) = -\frac{k_0}{2\pi\eta_0} \int_{-\infty}^{\infty} \tilde{G}(k_x)\, \mathrm{e}^{-jk_x x}\, \mathrm{d}k_x \tag{2.39}$$

$$\tilde{G}(k_x) = \frac{1}{\tilde{D}(k_x)}, \quad \tilde{D}(k_x) = \frac{k_{z2}}{\varepsilon_2} + \frac{k_{z1}}{\varepsilon_1}, \quad k_{zi} = \sqrt{k_i^2 - k_x^2}, \quad \text{for} \quad i = 1, 2 \tag{2.40}$$

where $\tilde{G}(k_x)$ is the Green function in terms of the spectral variable k_x, ε_1 is the dielectric constant of the upper half-space and ε_2, containing a negative real part,

is the dielectric constant of the metal lower half-space. Assuming the permeabilities of the two regions are the same as that of free-space, the location of the surface wave pole k_{xp}, which can be found analytically by setting $\tilde{D}(k_x)$ to zero, is $k_{xp} = k_0\sqrt{\varepsilon_1\varepsilon_2/(\varepsilon_1+\varepsilon_2)}$ as predicted by (2.6). The corresponding residue is:

$$R_p = \frac{1}{\tilde{D}'(k_{xp})}, \qquad \tilde{D}'(k_x) = -k_x\left(\frac{1}{\varepsilon_2 k_{z2}} + \frac{1}{\varepsilon_1 k_{z1}}\right). \tag{2.41}$$

The spectral domain Green function $\tilde{G}(k_x)$ has branch points at k_1 and k_2, respectively the wavenumbers of air and the metal at optical frequencies. Air is lossless, so k_1 is real and therefore lies on the real axis, as shown in Fig. 2.14, whereas metal has significant loss, so k_2 is far removed from and below the real axis. The field can be obtained by integrating in the k_x-plane along the real axis, path C_0, from minus to plus infinity by means of the definitions $\mathrm{Im}(k_{z1}) < 0$ and $\mathrm{Im}(k_{z2}) < 0$ on the top sheet of the complex plane for both branch cuts. Integration on the real axis can be improved upon by deforming the integration path vertically so that C_0 is replaced by C_1 and C_2, paths of steepest descent in the k_x-plane. An important consequence of the optical properties of a noble metal is that the position of the pole k_{xp} is such that it is captured by the vertical path deformation. Had the real part of the dielectric constant of metal been positive, the pole would reside to the left of the integration path below k_1 and therefore, it would not be captured by the integration path. It has been shown that poles captured by the steepest descent path are physical waves, in this case a surface plasmon wave, whereas poles not captured are not actual waves although they do influence the behavior of the field through their proximity to the integration path [69]. The contribution from the integral C_2 around the branch point at k_2 is negligible because the branch point lies well below the real axis so exponential

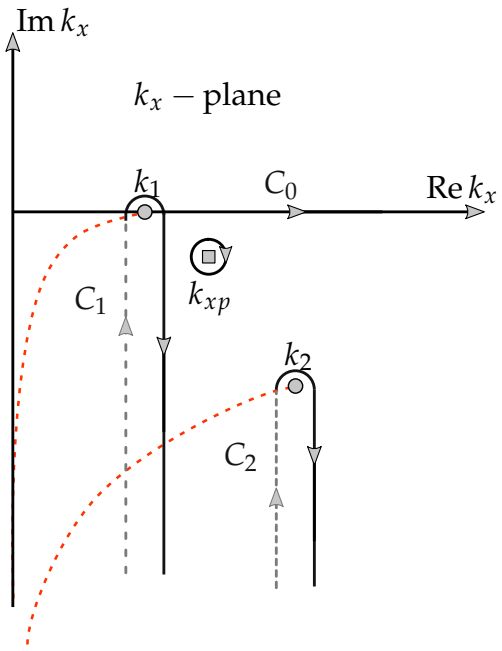


Figure 2.14: Complex plane integration path around branch cuts connecting the branch points k_1 and k_2, the wavenumbers for air and metal respectively, and the SPP pole k_{xp}.

decay due to the imaginary part of k_x is significant. The spatial domain Green function can now be expressed in terms of its pole and branch point components as,

$$G(x) = \underbrace{\int_{C_1} \tilde{G}(k_x)\, \mathrm{e}^{-jk_x x}\, \mathrm{d}k_x}_{\text{Composite Wave}} + \underbrace{\int_{C_2} \tilde{G}(k_x)\, \mathrm{e}^{-jk_x x}\, \mathrm{d}k_x}_{\text{Negligible}} - \underbrace{2\pi j R_p\, \mathrm{e}^{-jk_{xp} x}}_{\text{SPP Wave}}. \tag{2.42}$$

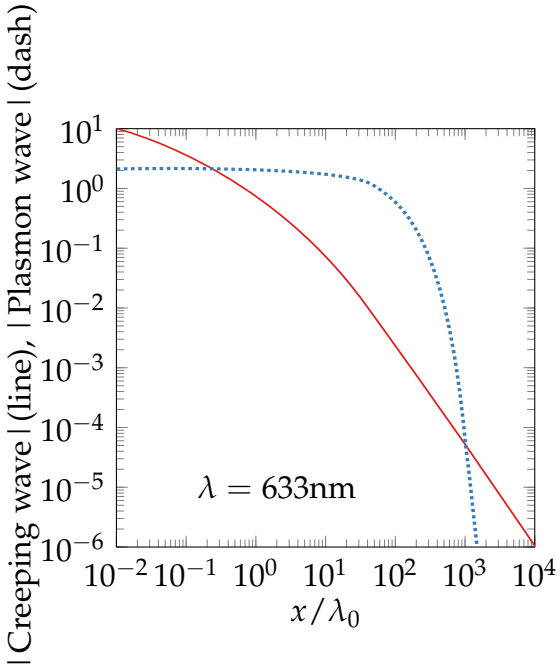
A careful analytical study of the composite wave shows that subtracting the surface plasmon polariton pole from the C_1 integrand yields a negligible result. However, when the pole is added back the resulting analytical expression yields

the following limiting case forms [17],

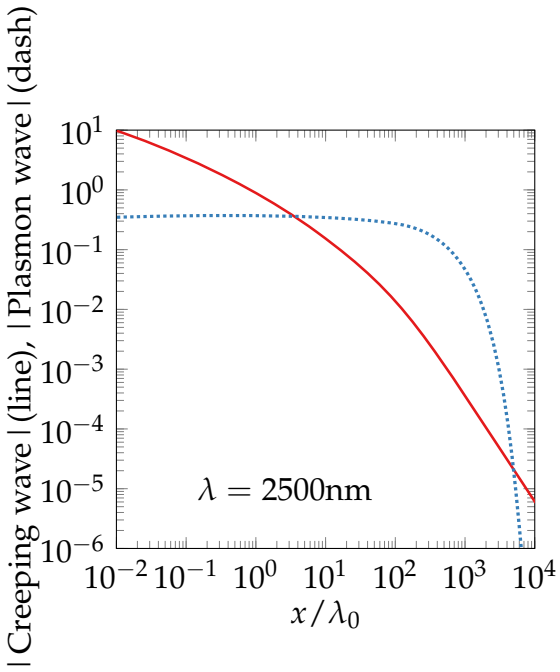
$$I_p(x) \sim \begin{cases} x^{-1/2}, & \text{for small and moderate distance } x \\ x^{-3/2}, & \text{for sufficiently large distance } x \end{cases} \tag{2.43}$$

At small or moderate distances from the aperture (magnetic line source) the magnetic field has a $x^{-1/2}$ behavior, the same as the small argument behavior of a Hankel function for a line source in free-space, is often referred to as the 'space wave'. At large distances, the $x^{-3/2}$ decay is typical of a lateral wave behavior.

Numerical results displaying the composite wave (solid line) and surface plasmon polariton wave (dashed line) where the metal is silver and the line source is radiating at 633 nm and 2500 nm are shown in Fig. 2.15 below. As predicted the dominant behavior of the magnetic field near the aperture is the $x^{-1/2}$ space wave. However at greater distances the plasmon wave becomes dominant, and in Fig. 2.15a close to 500 wavelengths where the plasmon wave tends to rapidly decay, the lateral wave behavior $x^{-3/2}$ dominates. Fig. 2.15b shows that at 2500 nm, the surface wave amplitude is reduced due to losses that enter the numerical calculation through the dielectric function given in Fig. 2.5. Eventually, at frequencies below the near infrared range, the metal begins to behave more like a perfect conductor and the space wave completely dominates the behavior of the field. This exercise confirms the existence of the plasmon wave and displays its relationship to other wave phenomena produced in the aperture of a subwavelength slot and on the air-metal interface. More important from an antenna engineering standpoint is that energy carried by the plasmon wave through the slot does radiate into the space wave above the slot. The question, "Can a plasmon wave pass through a slot?" is answered by the analysis at the beginning of



(a)



(b)

Figure 2.15: Logarithmic scale comparison of composite wave and surface plasmon polariton wave normalized amplitude for an air-silver interface as a function of distance from the source at (a) 633 nm (b) 2500 nm

this section and in an excellent review covering analytical methods used to determine the properties of subwavelength apertures [70]. Because a surface plasmon wave is tightly bound to the metal surface, decaying to $1/e$ at distances far less than a wavelength above the metal walls, its expression in a waveguide is not governed by the typical waveguide boundary conditions. Plasmon waves in the two walls of the slot, although not entirely independent of one another, are not subjected to cutoff conditions, and can therefore carry much more energy than a waveguide evanescent mode or even an infinite sum of evanescent modes. The conclusion that can be drawn from this analysis is that at optical wavelengths, a significant amount of plasmon energy can be transported through a metallic slot with a sub-wavelength width. Because the slot is sub-wavelength in width, it is possible to form an array of plasmonic radiators to focus an intense beam to a sub-wavelength line or spot.

2.5.8 Bowtie Aperture Antenna

Many aperture antennas are a simple Babinet equivalent of their metal counterpart. The aperture bowtie, also known as the diabolo antenna, is one such design. The aperture bowtie shown in Fig. 2.1c differs from the well-known metal bowtie nanoantenna in that the opposing pair of aperture triangles are connected through an aperture extension of the facing triangular tips. Thus instead of a large charge density accumulating between the triangles across the air gap, a high optical current density develops within the gap but between the narrowly spaced metal sides of the gap. A high intensity enhanced field component is achieved in the gap connecting the apposal triangular arms by polarizing the incident light parallel to the narrow gap connecting the arms. Polarizing the incident field across the narrow dimension of the gap will produce little or no

field enhancement. One of the interesting properties of the aperture bowtie is that the magnetic field rather than the electric field is enhanced in the gap. By analogy, the metal bowtie behaves more like an electric dipole and the aperture bowtie more like a magnetic dipole. Numerical simulations have shown a 2900-fold enhancement of the magnetic field at a wavelength of 2540 nm, confined to a $40 \times 40\text{nm}^2$ region near the center of the nanoantenna.

Although in general aperture antenna enhancement is not as high as can be achieved with a similar metallic structure, apertures are easier to make and have the advantage of providing a visual objective. The aperture bowtie has been shown to have sufficient enhancement and confinement to be effective in nanolithography applications [71]. The effectiveness of the aperture bowtie nanoantenna in nanolithographic applications is attributed to the fact that the magnetic field created in the aperture gap above the photoresist mask enters the mask in the perpendicular direction, easily penetrating into the metal, thereby leading to noticeable dissipation effects [72].

Optical tweezers are instruments that use light to move microscopic particles. Most often this requires a high powered focused laser beam to provide the necessary force to trap and move the particle. Optical tweezers are often needed in biological applications and are expected to have important applications in the construction of nano-machines. A variant on the bowtie design has been shown to create an optical vortex in the electromagnetic field [73]. The unusual field pattern is such that a small particle moving away from the center of the vortex experiences increasingly higher forces pointing back towards the center. The aperture bowtie provides the user the ability to see the particle while manipulations are carried out.

2.5.9 Concentric Rings

Perhaps the most useful aperture antenna enhancement design is a concentric ring structure. In cross-section the ring has the appearance of a shallow trough etched into the metallic surface. Concentric rings have been incorporated around a variety of aperture designs with the goal of concentrating the transmitted signal into a narrow beam or to gain maximized reception. A single radiating aperture must be large in terms of wavelengths and contain a flat planar phase front in order to concentrate the signal. A common example at microwave frequencies is a parabolic dish antenna which in practice will have an aperture size of at least 10 wavelengths. However, a typical single sub-wavelength aperture antenna will have a large beamwidth. In order to provide directionality, an optical aperture antenna can be modified by adding the concentric ring structure which in essence provides a Bragg reflection condition for plasmon current waves traveling out from center. Another way of looking at the concentric ring design is that it provides an external impedance match condition for the source driven aperture.

Again, because metal at optical frequencies has a behavior similar to that of a dielectric, at least in terms of boundary conditions, numerical solutions provide the only route to design criteria for concentric rings. The number of grooves, groove width, groove depth, periodicity of the grating, hole diameter, and the distance between the first groove and the driven aperture all play a role and are strongly interlinked in the design. However, a basic optimization criteria for bull's eye structures has been formulated based on extensive numerical studies. What follows is repeated from an excellent discussion on the subject, which can be found in more detail in a paper by [74]: "The first item to define is the desired resonance wavelength which in turn determines the period of the structure. The

choice of the hole diameter will then be determined by whether one would like optimal efficiency as normalized to hole area or highest absolute transmission. For the former, the diameter should be about half the period, but for the latter the aperture size can be increased. It should be noted that as the hole size increases relative to the period, the spectrum will eventually broaden which is a trade-off. The groove width should also be around half the period with a depth to width ratio at 0.4. The number of grooves should be just enough to reach saturation, typically around 6 to 10 grooves depending on the geometrical parameters."

2.5.10 Arrays

Many other specialized aperture designs have been reported in the literature, although design information is scarce. Usually the designer can refer to the current microwave antenna literature as a starting point for a numerical parameter investigation to determine that proper dimensions for an optical antenna design. External to the antenna the array factor and standard microwave array methods will typically be applicable to optical antenna array design, assuming the aperture field information is known. That is, a computer code must be used to determine the near field and a suitable array factor can be applied to the aperture near field information to find the far field. The array elements can be square or circular or any one of the antennas discussed above. Rectangular shaped apertures are of particular interest since they are the Babinet equivalent of a dipole antenna. The half-wavelength resonance condition applies to the slot antenna of high conductivity and a half-wave slot leads to a resonantly enhanced and aperture confined electric field [25].

2.6 Applications of Optical Antenna

2.6.1 SNOM

Scanning near-field optical microscopy (SNOM), an important technique for visualizing biological systems, involves obtaining high resolution topographic and optical images with a cone shaped probe that scales down to the nano level at the tip. The need for high spatial, temporal and resolving power with sub-wavelength resolution was the driving factor in a move to the fiber aperture probe, essentially a tapered optical fiber with a nanoscale tip. Evanescent or non-propagating fields that exist only near the surface of the object carry the high frequency spatial information about the object and have intensities that drop off exponentially with distance from the object. Because of this, the detector must be placed very close to the sample in the near field zone, typically a few nanometers. However, the aperture probe suffers from diffraction of light at the tip, which reduces its imaging resolution. Present research is centered on finding an antenna design that reduces the fiber beamwidth, thereby improving resolution. Although fabrication of tips true to the design criteria and surface smoothness are significant challenges to be overcome, a number of antenna structures such as the bowtie, Yagi and several plasmonic nanosphere probes have been tried out showing successful improvement in many cases.

2.6.2 Photon Emitters and Optical Fluorescence

Single photon emission from a quantum emitter generates a stream of photons. Single photons can be used in a number of technologies such as Computed Tomography (CT), which is an imaging technique able to provide 3D information. A single quantum emitter positioned inside the subwavelength size feed gap of an optical antenna couples to the antenna, radiating in such a way as to

be considered to be an impedance matching device. The density of the emitted photons thereby increases, improving the tomographic image.

Solid state light-emitting devices are expected to eventually replace fluorescent tubes as illumination sources. Quantum dot nanocrystals are very promising for light-emitting sources. However, their light-emission efficiencies are still substantially lower than those of fluorescent tubes. It has been observed that quantum emitters in close proximity to a plasmonic producing metal surface causes plasmonic fluctuations of the free electron gas. The associated currents radiate, in many cases with a significant increase in intensity. Antenna structures such as the bowtie have been designed, optimized in size, shape, and material properties to increase the radiation efficiency of a nanoscopic optical source. However, for visual appeal, sources with greater optical bandwidth are needed [18, 75].

2.6.3 Raman Spectroscopy

Molecules vibrate, rotate and translate in a number of ways when exposed to an electromagnetic field. Raman spectroscopy is an optical technique used to detect the vibrational modes. It relies upon detection of a very low intensity portion of the scattered wave spectrum called inelastic or Raman scattering. Because molecules of a particular type, such as anthrax or mold aflatoxins, have a certain Raman spectra they can be identified by this method. However, because the Raman spectra is very small, very intense laser light is required in order to move the Raman signal above the noise level. Plasmonic nanoantennas create highly enhanced local fields when pumped resonantly, leading to increased Raman scattering [76].

2.6.4 Communication with Nanocircuitry

As circuit chips become smaller and processor speeds increase, wire leads to the circuitry become less practical and less efficient. However, nanoantennas can overcome these drawbacks, with more efficient and on chip created optical antennas communicating over wireless links to external circuitry. Cross-talk and the many problems with computer interconnects can be eliminated. One of the properties of on-surface optical antennas is that most of the radiated energy travels downward into the chip substrate. This can become an advantage where the construction of a via through the nano chip is impractical. A proper nanoantenna design would allow the circuitry to rapidly communicate back and forth directly through the substrate.

In this section, an overview of the field of optical antennas has been presented along with some of the basic theory of plasmonics, the underpinning concept governing electromagnetic wave behavior at optical frequencies. It was shown that surface plasmon polaritons lead to a set of requirements and restrictions placed on the design of antennas constructed out of noble metals at optical and near infrared frequencies. An unusual aspect of plasmonic waves in metals is their subwavelength behavior, an advantage when detecting small particles but it can cause added difficulty when impedance matching an optical antenna to free-space. It was shown that the surface plasmons explain the unusual amount of radiation coupled through a subwavelength aperture and that as a consequence aperture optical antennas are remarkably efficient, behaving in a manner similar to their Babinet equivalent counterpart. Spectroscopy, disease and toxin sensors, wireless communication with nano circuitry and the creation of nano circuits using sub wavelength lithography are some of the early beneficiaries of this new

technology. The field of nano plasmonics offers great opportunities for optical frequency circuit and antenna engineers with the future easily as bright as it is for the modern day wireless engineer.

Layered media are the basis of modern semiconductor integrated circuits that have been the subject of numerous theoretical studies designed to gain an understanding of how the electromagnetic fields behave in an environment in which the source or receiver is embedded in layers with different thicknesses and material properties. Most commonly, the fields are found in terms of the dyadic Green functions (DGFs) of the environment, which defines the vector field distribution due to a vector point source [77]. The surface fields are expressed in terms of a current $\mathbf{J}$ distributed in a region defined by a surface, r', by the inner product,

$$\mathbf{E} = \int_{r'} \underline{\underline{\mathbf{G}}}^{\mathrm{EJ}}(\mathbf{r}|\mathbf{r}')\mathbf{J}(\mathbf{r}')\,\mathrm{d}r', \tag{3.1a}$$

$$\mathbf{H} = \int_{r'} \underline{\underline{\mathbf{G}}}^{\mathrm{HJ}}(\mathbf{r}|\mathbf{r}')\mathbf{J}(\mathbf{r}')\,\mathrm{d}r'. \tag{3.1b}$$

where, $\underline{\underline{\mathbf{G}}}^{\mathrm{EJ}}$ and $\underline{\underline{\mathbf{G}}}^{\mathrm{HJ}}$ are the spatial domain DGF's due to an electric current source located at r', that define the electric and magnetic fields at a point r respectively. In the presence of magnetic sources, (3.1a) and (3.1b) can be augmented using the superposition principle, by adding an inner product containing magnetic DGFs, $\underline{\underline{\mathbf{G}}}^{\mathrm{EM}}$ and $\underline{\underline{\mathbf{G}}}^{\mathrm{HM}}$.

3.1 Theory

Here we follow the well-established approach of field computation for planar, layered media [2, 8], in which, an equivalent transmission line (TL) network is first set up for the structure, and then TL based GFs are computed.

3.1.1 Transmission Line Representation of Maxwell's Equations

As shown in Fig. 3.3a, it is assumed that the structure is unbounded in the lateral direction, and excited only by an electric source. The electric and magnetic fields are given by the Maxwell's equations in frequency-domain,

$$\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}, \tag{3.2a}$$

$$\nabla \times \mathbf{H} = j\omega\varepsilon\mathbf{E} + \mathbf{J}. \tag{3.2b}$$

For boundary-value problems displaying symmetry along the z direction, it is desirable to decompose the ∇ operator into two components, one longitudinal d/dz and the other a transverse (to z) operator, $\nabla_{\mathbf{t}}$ [78, p. 64]. By taking the Fourier transform,

$$\mathcal{F}[f(\mathbf{k}_\rho, z)] \equiv \tilde{F}(\mathbf{k}_\rho, z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\boldsymbol{\rho}, z)\mathrm{e}^{-j\mathbf{k}_\rho\cdot\boldsymbol{\rho}}\,\mathrm{d}x\,\mathrm{d}y \tag{3.3}$$

the field computation is considerably simplified by switching to the spectral frequency domain $\mathbf{k}_\rho$, which reduces the complexity of the vector differential operator, ∇ to $-jk_x\hat{\mathbf{x}} - jk_y\hat{\mathbf{y}} + d/dz\,\hat{\mathbf{z}}$, which contains a derivative term only in z-direction. In (3.3), the cylindrical coordinate vectors are expressed as,

$$\boldsymbol{\rho} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}, \quad \text{and} \quad \mathbf{k}_{\boldsymbol{\rho}} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}}, \tag{3.4}$$

and the notation $\sim$ above the capital-letter terms indicates the Fourier transform with respect to the transverse coordinates and from here on, will be used to denote the spectral domain quantities.

As stated earlier, it is advantageous to separate the fields in transverse and

longitudinal coordinates because, as we shall see shortly, the longitudinal part of the field can be completely expressed in terms of the transverse component. Applying the Fourier transform (3.3) on the Maxwell's equations (3.2), we obtain:

$$\left(-j\mathbf{k}_\rho + \hat{\mathbf{z}}\frac{\mathrm{d}}{\mathrm{d}z}\right) \times (\tilde{\mathbf{E}}_t + \tilde{\mathbf{E}}_z) = -j\omega\mu(\tilde{\mathbf{H}}_t + \tilde{\mathbf{H}}_z), \tag{3.5a}$$

$$\left(-j\mathbf{k}_\rho + \hat{\mathbf{z}}\frac{\mathrm{d}}{\mathrm{d}z}\right) \times (\tilde{\mathbf{H}}_t + \tilde{\mathbf{H}}_z) = j\omega\varepsilon(\tilde{\mathbf{E}}_t + \tilde{\mathbf{E}}_z) - (\tilde{\mathbf{J}}_t + \tilde{\mathbf{J}}_z). \tag{3.5b}$$

The transverse and longitudinal components of the magnetic field can be separately expressed in (3.5a) as,

$$-j\mathbf{k}_\rho \times \tilde{\mathbf{E}}_z + \frac{\mathrm{d}}{\mathrm{d}z}\hat{\mathbf{z}} \times \tilde{\mathbf{E}}_t = -j\omega\mu\tilde{\mathbf{H}}_t, \tag{3.6a}$$

$$-j\mathbf{k}_\rho \times \tilde{\mathbf{E}}_t = -j\omega\mu\tilde{\mathbf{H}}_z. \tag{3.6b}$$

Using the vector product property [79, p. 117],

$$\mathbf{A} \times \mathbf{B} = \mathbf{A} \cdot (\mathbf{B} \times \hat{\mathbf{n}})\,\hat{\mathbf{n}}, \tag{3.7}$$

where the unit vector $\hat{\mathbf{n}}$ is normal to the plane containing vectors $\mathbf{A}$ and $\mathbf{B}$. A scalar form of the longitudinal component of the electric field is obtained by applying (3.7) on (3.6b),

$$-j\mathbf{k}_\rho \cdot (\tilde{\mathbf{E}}_t \times \hat{\mathbf{z}})\,\hat{\mathbf{z}} = -j\omega\mu\tilde{\mathbf{H}}_z \tag{3.8}$$

which can be written in the scalar form,

$$-j\tilde{H}_z = \frac{-j}{\omega\mu}\mathbf{k}_\rho \cdot (\tilde{\mathbf{E}}_t \times \hat{\mathbf{z}}). \tag{3.9}$$

Now taking the vector product with unit vector $\hat{\mathbf{z}}$ on both sides of (3.6a), the transverse electric field component is expressed as:

$$\begin{aligned} \frac{\mathrm{d}\tilde{\mathbf{E}}_t}{\mathrm{d}z} &= -j(\mathbf{k}_\rho \times \tilde{\mathbf{E}}_z) \times \hat{\mathbf{z}} - j\omega\mu\tilde{\mathbf{H}}_t \times \hat{\mathbf{z}} \\ &= -j\mathbf{k}_\rho \tilde{E}_z - j\omega\mu\tilde{\mathbf{H}}_t \times \hat{\mathbf{z}} \end{aligned} \tag{3.10}$$

where the BAC-CAB vector triple product identity, $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} \equiv \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ has been applied.

Following a similar procedure starting from (3.5b), we obtain the transverse component of magnetic field, along with the scalar longitudinal component of electric field:

$$\begin{aligned} \frac{\mathrm{d}\tilde{\mathbf{H}}_t}{\mathrm{d}z} &= -j(\mathbf{k}_\rho \times \tilde{\mathbf{H}}_z) \times \hat{\mathbf{z}} + j\omega\varepsilon\tilde{\mathbf{E}}_t \times \hat{\mathbf{z}} + \tilde{\mathbf{J}}_t \times \hat{\mathbf{z}} \\ &= -j\mathbf{k}_\rho \tilde{H}_z + j\omega\varepsilon\tilde{\mathbf{E}}_t \times \hat{\mathbf{z}} + \tilde{\mathbf{J}}_t \times \hat{\mathbf{z}}, \end{aligned} \tag{3.11}$$

and,

$$-j\omega\varepsilon\tilde{E}_z = j\mathbf{k}_\rho \cdot (\tilde{\mathbf{H}}_t \times \hat{\mathbf{z}}) + \tilde{J}_z. \tag{3.12}$$

By substituting (3.12) in (3.10), we get the transverse component of electric field,

$$\frac{\mathrm{d}\tilde{\mathbf{E}}_t}{\mathrm{d}z} = \frac{1}{j\omega\varepsilon}\left(k^2 - \mathbf{k}_\rho\mathbf{k}_\rho\cdot\right)(\tilde{\mathbf{H}}_t \times \hat{\mathbf{z}}) + \mathbf{k}_\rho\frac{\tilde{J}_z}{\omega\varepsilon}. \tag{3.13}$$

Similarly, from (3.9) and (3.11), the transverse component of magnetic field,

$$\frac{\mathrm{d}\tilde{\mathbf{H}}_t}{\mathrm{d}z} = \frac{1}{j\omega\mu}\left(k^2 - \mathbf{k}_\rho \mathbf{k}_\rho \cdot\right)(\hat{\mathbf{z}} \times \tilde{\mathbf{E}}_t) + \tilde{\mathbf{J}}_t \times \hat{\mathbf{z}} \tag{3.14}$$

where $k = \omega\sqrt{\mu\varepsilon}$ in (3.13) and (3.14) is the medium wavenumber.

The fields in (3.13) and (3.14) for arbitrarily aligned sources lie in the plane of a spectral coordinate system as illustrated in Fig. 3.1, where the arrowheads in color correspond to spectral-domain quantities. A rotational transformation of the coordinate system such that the axes align with the vectors $\mathbf{k}_\rho, \hat{\mathbf{z}} \times \mathbf{k}_\rho$ [80], simplifies the procedure of finding the transmission line equivalent, which allows the TE and TM mode analysis to be made separately. The coordinate transformation can be expressed as:

$$\begin{bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \end{bmatrix} \tag{3.15}$$

where ψ is the angle between $\mathbf{k}_{\mathbf{æ}}$ and the positive x-axis. A transmission line analogue for the spectral fields, expressed in terms of modal voltages and currents can therefore, be written as [81, 8],

$$\begin{bmatrix} \tilde{\mathbf{E}}_t \\ \tilde{\mathbf{H}}_t \end{bmatrix} = \begin{bmatrix} V^{\mathrm{TM}} & V^{\mathrm{TE}} \\ -I^{\mathrm{TE}} & I^{\mathrm{TM}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{v}} \end{bmatrix}. \tag{3.16}$$

Using the results of (3.16) in (3.13) and noting that $\hat{\mathbf{u}} = \mathbf{k}_\rho / k_\rho$, we get,

$$\frac{\mathrm{d}\left(\hat{\mathbf{u}}\, V^{\mathrm{TM}} + \hat{\mathbf{v}}\, V^{\mathrm{TE}}\right)}{\mathrm{d}z} = \frac{1}{j\omega\varepsilon}\left(k^2 - \mathbf{k}_\rho\, \mathbf{k}_\rho \cdot\right)(\hat{\mathbf{u}}\, I^{\mathrm{TM}} + \hat{\mathbf{v}}\, I^{\mathrm{TE}}) + \hat{\mathbf{u}}\, \frac{k_\rho}{\omega\varepsilon} \tilde{J}_z \tag{3.17}$$

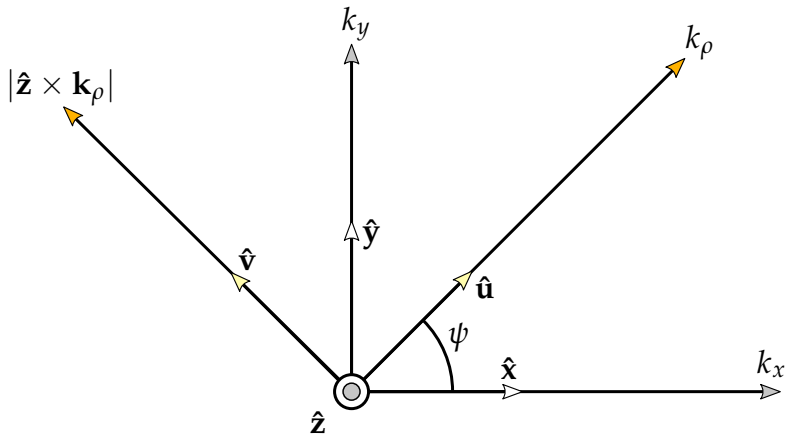


Figure 3.1: Coordinate system transformation in the spectral domain

By separating the $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ components, we obtain the TM and TE equivalent voltage equations respectively,

$$\frac{\mathrm{d}V^{\mathrm{TM}}}{\mathrm{d}z} = \frac{1}{j\omega\varepsilon}\left(k^2 - k_\rho^2\right) I^{\mathrm{TM}} + \frac{k_\rho}{\omega\varepsilon}\tilde{J}_z, \tag{3.18a}$$

$$\frac{\mathrm{d}V^{\mathrm{TE}}}{\mathrm{d}z} = \frac{k^2}{j\omega\varepsilon} I^{\mathrm{TE}}. \tag{3.18b}$$

Similarly, from (3.16) and (3.14), the equivalent current equations can be written as:

$$\frac{\mathrm{d}I^{\mathrm{TM}}}{\mathrm{d}z} = \frac{k^2}{j\omega\mu} V^{\mathrm{TM}} - \tilde{J}_u, \tag{3.19a}$$

$$\frac{\mathrm{d}I^{\mathrm{TE}}}{\mathrm{d}z} = \frac{-1}{j\omega\mu}\left(k^2 - k_\rho^2\right) V^{\mathrm{TE}} + \tilde{J}_v. \tag{3.19b}$$

Equations (3.18)-(3.19) can be conveniently written in a compact form as a set of

telegrapher's equations [78, p. 190]:

$$\frac{\mathrm{d}V^{\alpha}}{\mathrm{d}z} = -jk_z Z^{\alpha} I^{\alpha} + v^{\alpha} \tag{3.20a}$$

$$\frac{\mathrm{d}I^{\alpha}}{\mathrm{d}z} = -jk_z Y^{\alpha} V^{\alpha} + i^{\alpha} \tag{3.20b}$$

where α is either TE or TM, the propagation constant in the transverse direction is $k_z = \pm\sqrt{k^2 - k_\rho^2}$, for which the sign must be chosen so that the fields decay away from the source. The modal impedances in (3.20) are,

$$Z^{\mathrm{TM}} = \frac{1}{Y^{\mathrm{TM}}} = \frac{k_z}{\omega\varepsilon}, \tag{3.21a}$$

$$Z^{\mathrm{TE}} = \frac{1}{Y^{\mathrm{TE}}} = \frac{\omega\mu}{k_z}. \tag{3.21b}$$

Using the expressions that relate the transverse electric and magnetic to the equivalent transmission line currents and voltages (3.16), and combining with the longitudinal field expressions (3.6b) and (3.12), we obtain the total fields in the spectral domain,

$$\begin{bmatrix} \tilde{\mathbf{E}}(\mathbf{k}_\rho, z) \\ \tilde{\mathbf{H}}(\mathbf{k}_\rho, z) \end{bmatrix} = \begin{bmatrix} V^{\mathrm{TM}} & V^{\mathrm{TE}} & -\frac{k_\rho}{\omega\varepsilon(z)} I^{\mathrm{TM}}(z) \\ -I^{\mathrm{TE}} & I^{\mathrm{TM}} & \frac{k_\rho}{\omega\mu} V^{\mathrm{TE}}(z) \end{bmatrix} \begin{bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{v}} \\ \hat{\mathbf{z}} \end{bmatrix} + \hat{\mathbf{z}} \begin{bmatrix} \frac{j}{\omega\varepsilon(z)} \tilde{\mathbf{J}}_z(\mathbf{k}_\rho, z) \\ 0 \end{bmatrix}, \tag{3.22}$$

where $\varepsilon(z)$ may vary from one layer to another.

Assuming only electric sources existing in space, the corresponding TL sources, v^{α} and i^{α}, defined in (3.20), are illustrated in 3.2. A horizontally oriented (x-directed) electric dipole is represented by a current source in an equivalent TM

transmission line network. Likewise, the equivalent configuration of a vertical (y-directed) electric dipole is a TE network with a current source. A z-directed dipole corresponds to voltage source in a TM transmission line. For an arbitrarily directed source, the equivalent TL model consists of a superposition of the three representations.

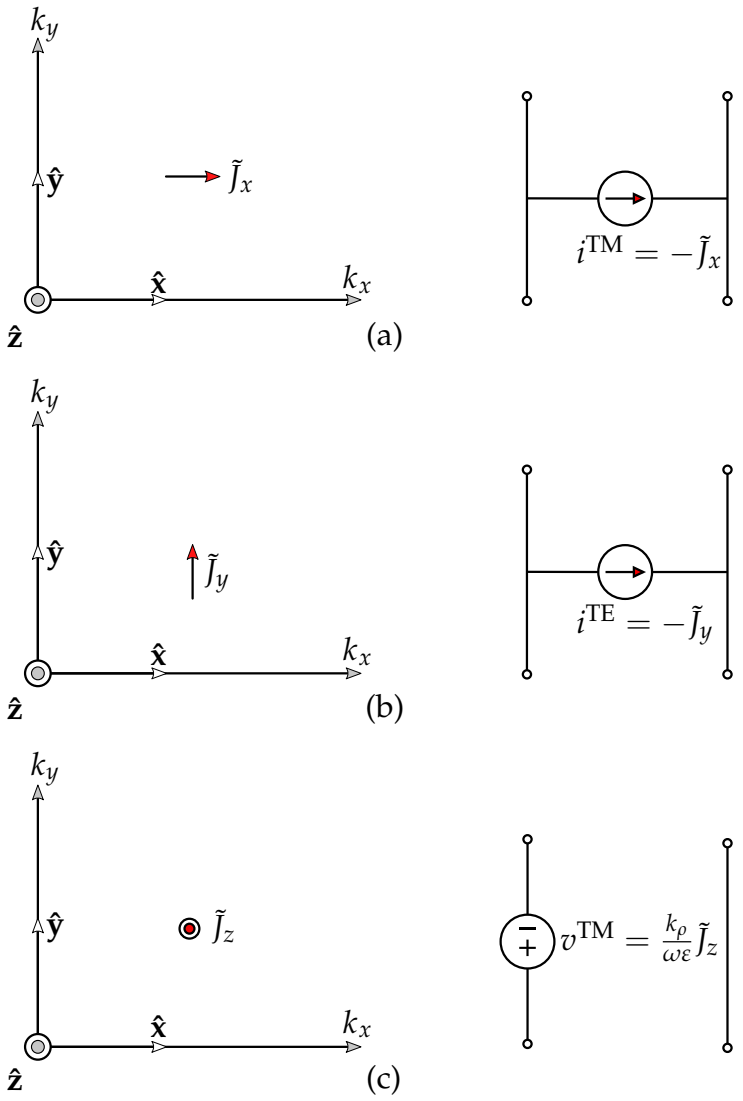


Figure 3.2: Electric source representation in a transmission line network

3.1.2 Green Functions for the TL Equations

To obtain the transmission line voltages and currents that define the spectral fields in (3.22), we introduce the one-dimensional transmission line Green functions (TLGFs) that are analogous to the spatial domain DGFs in (3.1). Following [78, 2], we define $V_i(z, z')$ and $I_i(z, z')$ as the voltage and current respectively, at a point z along the transmission line due to a unit-strength current source located at z'. Similarly, $V_v(z, z')$ and $I_v(z, z')$ are the respective voltage and current due to a unit-strength voltage source. Analogous to (3.1) which is augmented by magnetic sources, we write:

$$V(z) = \int_{z'} \left[V_i(z, z') i(z') + V_v(z, z') v(z') \right] \mathrm{d}z', \tag{3.23a}$$

$$I(z) = \int_{z'} \left[I_i(z, z') i(z') + I_v(z, z') v(z') \right] \mathrm{d}z'. \tag{3.23b}$$

The telegrapher's equations (3.20), are rewritten for a voltage excited line as:

$$\frac{\mathrm{d}V_v}{\mathrm{d}z} = -jk_z Z_v I_v + \delta(z - z') \tag{3.24a}$$

$$\frac{\mathrm{d}I_v}{\mathrm{d}z} = -jk_z Y_v V_v, \tag{3.24b}$$

and for a current excited line, we obtain:

$$\frac{\mathrm{d}V_i}{\mathrm{d}z} = -jk_z Z_i I_i \tag{3.25a}$$

$$\frac{\mathrm{d}I_i}{\mathrm{d}z} = -jk_z Y_i V_i + \delta(z - z') \tag{3.25b}$$

3.1.3 Spectral Domain Dyadic Green Functions

By substituting (3.24) and (3.25) into (3.22) and referring to Fig. 3.2, we obtain the spectral-domain versions of DGFs [8]:

$$\begin{aligned}\underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{EJ}}(\mathbf{k}_\rho, z|z') = & -\hat{\mathbf{u}}\hat{\mathbf{u}}V_i^{\mathrm{TM}} - \hat{\mathbf{v}}\hat{\mathbf{v}}V_i^{\mathrm{TE}} + \hat{\mathbf{z}}\hat{\mathbf{u}}\frac{k_\rho}{\omega\varepsilon(z)}I_i^{\mathrm{TM}} \\ & + \hat{\mathbf{u}}\hat{\mathbf{z}}\frac{k_\rho}{\omega\varepsilon(z')}V_v^{\mathrm{TM}} + \hat{\mathbf{z}}\hat{\mathbf{z}}\frac{1}{j\omega\varepsilon(z')}\left[\frac{k_\rho^2}{j\omega\varepsilon(z)}I_v^{\mathrm{TM}} - \delta(z-z')\right], \quad (3.26a)\end{aligned}$$

$$\underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{HJ}}(\mathbf{k}_\rho, z|z') = \hat{\mathbf{u}}\hat{\mathbf{v}}I_i^{\mathrm{TE}} - \hat{\mathbf{v}}\hat{\mathbf{u}}I_i^{\mathrm{TE}} - \hat{\mathbf{z}}\hat{\mathbf{v}}\frac{k_\rho}{\omega\mu}V_i^{\mathrm{TE}} + \hat{\mathbf{v}}\hat{\mathbf{z}}\frac{k_\rho}{\omega\varepsilon(z')}I_v^{\mathrm{TM}}. \quad (3.26b)$$

3.1.4 Spatial Domain Dyadic Green Functions

By taking an inverse Fourier transform (3.27),

$$\mathcal{F}^{-1}[\tilde{F}(\mathbf{k}_\rho, z)] \equiv f(\boldsymbol{\rho}, z) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tilde{F}(\mathbf{k}_\rho, z)\mathrm{e}^{j\mathbf{k}_\rho\cdot\boldsymbol{\rho}}\,\mathrm{d}k_x\,\mathrm{d}k_y. \quad (3.27)$$

we find the spatial domain analogues of the spectral domain DGFs, which are defined in (3.26). In case of rotational symmetry which implies that $\tilde{F}$ only depends on one spectral co-ordinate k_ρ, the double integral in (3.27) can be simplified to an integral of only one variable. Using (3.4) and the coordinate transformation shown in Fig. 3.1 where,

$$k_x = k_\rho\cos\psi \qquad \text{and} \qquad k_y = k_\rho\sin\psi, \quad (3.28)$$

we rewrite (3.27) in cylindrical coordinates,

$$\mathcal{F}^{-1}[\tilde{F}(\mathbf{k}_\rho, z)] \equiv f(\boldsymbol{\rho}, z) = \frac{1}{(2\pi)^2} \int_0^\infty \int_0^{2\pi} \tilde{F}(\mathbf{k}_\rho, z) \mathrm{e}^{j\mathbf{k}_\rho \cdot \boldsymbol{\rho}} k_\rho \, \mathrm{d}k_\rho \, \mathrm{d}\psi. \tag{3.29}$$

Applying the Fourier-Bessel transform (FBT) to (3.29) which states that,

$$\frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{j\mathbf{k}_\rho \cdot \boldsymbol{\rho}} \, \mathrm{d}\psi = J_0(k_\rho \rho), \tag{3.30}$$

where $J_0(\cdot)$ is the Bessel function of zero order we obtain the Sommerfeld integral (SI), $\mathcal{S}_0\{\cdot\}$:

$$f(\rho) = \mathcal{S}_0\left\{\tilde{F}\right\} \equiv \frac{1}{2\pi} \int_0^\infty J_0(k_\rho \rho) \tilde{F}(k_\rho) k_\rho \, \mathrm{d}k_\rho. \tag{3.31}$$

In some cases, the Bessel function may be up to an order of 2, therefore, generalized expression for SI is,

$$\mathcal{S}_n\left\{\tilde{F}\right\} \equiv \frac{1}{2\pi} \int_0^\infty J_n(k_\rho \rho) \tilde{F}(k_\rho) k_\rho \, \mathrm{d}k_\rho. \tag{3.32}$$

Following the above procedure, we write the spatial domain DGFs as:

$$\underline{\underline{\mathbf{G}}}^{\kappa}(\boldsymbol{\rho}, z|z') = \mathcal{F}^{-1}\left[\underline{\underline{\tilde{\mathbf{G}}}}^{\kappa}(\mathbf{k}_\rho, z|z')\right] \equiv \mathcal{S}_n\left\{\underline{\underline{\tilde{\mathbf{G}}}}^{\kappa}(\mathbf{k}_\rho, z|z')\right\} \tag{3.33}$$

where the superscript κ denotes either 'EJ' or 'HJ'.

Starting with (3.26), we now discuss each step in detail to obtain the spatial domain versions of the DGFs. First, the spectral domain dyads are converted to their spatial domain counterparts using the relations listed in Table 3.1 which are obtained from Fig. 3.1 and (3.15). In general, a spatial domain DGF can be

Spectral domain dyad	Spatial domain dyad
$\hat{\mathbf{u}}\hat{\mathbf{u}}$	$\hat{\mathbf{x}}\hat{\mathbf{x}}\cos^2\psi + \hat{\mathbf{x}}\hat{\mathbf{y}}\cos\psi\,\sin\psi + \hat{\mathbf{y}}\hat{\mathbf{x}}\cos\psi\,\sin\psi + \hat{\mathbf{y}}\hat{\mathbf{y}}\sin^2\psi$
$\hat{\mathbf{v}}\hat{\mathbf{v}}$	$\hat{\mathbf{x}}\hat{\mathbf{x}}\sin^2\psi - \hat{\mathbf{x}}\hat{\mathbf{y}}\cos\psi\,\sin\psi - \hat{\mathbf{y}}\hat{\mathbf{x}}\cos\psi\,\sin\psi + \hat{\mathbf{y}}\hat{\mathbf{y}}\sin^2\psi$
$\hat{\mathbf{u}}\hat{\mathbf{v}}$	$-\hat{\mathbf{x}}\hat{\mathbf{x}}\cos\psi\,\sin\psi + \hat{\mathbf{x}}\hat{\mathbf{y}}\cos^2\psi - \hat{\mathbf{y}}\hat{\mathbf{x}}\sin^2\psi + \hat{\mathbf{y}}\hat{\mathbf{y}}\cos\psi\,\sin\psi$
$\hat{\mathbf{v}}\hat{\mathbf{u}}$	$-\hat{\mathbf{x}}\hat{\mathbf{x}}\cos\psi\,\sin\psi - \hat{\mathbf{x}}\hat{\mathbf{y}}\sin^2\psi + \hat{\mathbf{y}}\hat{\mathbf{x}}\cos^2\psi + \hat{\mathbf{y}}\hat{\mathbf{y}}\cos\psi\,\sin\psi$
$\hat{\mathbf{z}}\hat{\mathbf{u}}$	$\hat{\mathbf{z}}\hat{\mathbf{x}}\cos\psi + \hat{\mathbf{z}}\hat{\mathbf{y}}\sin\psi$
$\hat{\mathbf{u}}\hat{\mathbf{z}}$	$\hat{\mathbf{x}}\hat{\mathbf{z}}\cos\psi + \hat{\mathbf{y}}\hat{\mathbf{z}}\sin\psi$
$\hat{\mathbf{z}}\hat{\mathbf{v}}$	$-\hat{\mathbf{z}}\hat{\mathbf{x}}\sin\psi + \hat{\mathbf{z}}\hat{\mathbf{y}}\cos\psi$
$\hat{\mathbf{v}}\hat{\mathbf{z}}$	$-\hat{\mathbf{x}}\hat{\mathbf{z}}\sin\psi + \hat{\mathbf{y}}\hat{\mathbf{z}}\cos\psi$

Table 3.1: Conversion of spectral domain dyad to spatial domain

expressed in a matrix form:

$$\underline{\underline{\mathbf{G}}}^{\kappa} = \begin{bmatrix} G^{\kappa}_{xx} & G^{\kappa}_{xy} & G^{\kappa}_{xz} \\ G^{\kappa}_{yx} & G^{\kappa}_{yy} & G^{\kappa}_{yz} \\ G^{\kappa}_{zx} & G^{\kappa}_{zy} & G^{\kappa}_{zz} \end{bmatrix}, \tag{3.34}$$

As an example, we consider the spatial domain component, G^{EJ}_{xx} of the DGF,

which can be expressed as:

$$\begin{aligned} G_{xx}^{\mathrm{EJ}} &= \mathcal{F}^{-1}\left\{-\cos^2\psi V_i^{\mathrm{TM}} - \sin^2\psi V_i^{\mathrm{TE}}\right\} \\ &= -\mathcal{F}^{-1}\left\{\frac{1+\cos 2\psi}{2}V_i^{\mathrm{TM}} + \frac{1-\cos 2\psi}{2}V_i^{\mathrm{TE}}\right\} \\ &= -\frac{1}{2}\mathcal{S}_0\left\{V_i^{\mathrm{TM}} + V_i^{\mathrm{TE}}\right\} - \mathcal{F}^{-1}\left\{\frac{\cos 2\psi}{2}V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}\right\} \end{aligned} \tag{3.35}$$

Using the formula [2],

$$\mathcal{F}^{-1}\left\{\cos 2\psi\,\tilde{F}\right\} = -\cos 2\phi\,\mathcal{S}_2\left\{\tilde{F}\right\} \tag{3.36}$$

and the recurrence relation for Bessel-type functions [82],

$$J_{n+1}(z) = \frac{2n}{z}J_n(z) - J_{n-1}(z) \tag{3.37}$$

(3.35) is simplified to:

$$\begin{aligned} G_{xx}^{\mathrm{EJ}} = & -\frac{1}{2}\mathcal{S}_0\left\{V_i^{\mathrm{TM}}\right\} - \frac{1}{2}\mathcal{S}_0\left\{V_i^{\mathrm{TE}}\right\} + \frac{\cos 2\phi}{2}\mathcal{S}_2\left\{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}\right\} \\ = & -\frac{1}{2}\mathcal{S}_0\left\{V_i^{\mathrm{TM}}\right\} - \frac{1}{2}\mathcal{S}_0\left\{V_i^{\mathrm{TE}}\right\} + \\ & \frac{\cos 2\phi}{2}\left[\frac{2}{\rho}\mathcal{S}_1\left\{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}\right\} - \mathcal{S}_0\left\{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}\right\}\right]. \end{aligned} \tag{3.38}$$

After some algebraic manipulation and knowing the fact that Sommerfeld integrals obey the linearity principle, we obtain the final expression:

$$G_{xx}^{\mathrm{EJ}} = -\cos^2\phi\,\mathcal{S}_0\left\{V_i^{\mathrm{TM}}\right\} - \sin^2\phi\,\mathcal{S}_0\left\{V_i^{\mathrm{TE}}\right\} + \frac{\cos 2\phi}{2}\mathcal{S}_2\left\{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}\right\}. \tag{3.39}$$

Following similar procedure, the remaining components of both electric and

magnetic fields DGFs in the spatial domain can be obtained, which are listed in Table 3.2 [2].

GF	$\kappa = \mathrm{EJ}$
G_{xx}	$-\cos^2\psi\,\mathcal{S}_0\left\{V_i^{\mathrm{TM}}\right\} - \sin^2\psi\,\mathcal{S}_0\left\{V_i^{\mathrm{TE}}\right\} + \frac{\cos(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}}{k_\rho}\right\}$
G_{xy}	$-\frac{\sin(2\psi)}{2}\mathcal{S}_0\left\{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}\right\} + \frac{\sin(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}}{k_\rho}\right\}$
G_{xz}	$\frac{\eta_0}{jk_0\varepsilon(z')}\cos\psi\,\mathcal{S}_1\left\{k_\rho V_v^{\mathrm{TM}}\right\}$
G_{yx}	$-\frac{\sin(2\psi)}{2}\mathcal{S}_0\left\{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}\right\} + \frac{\sin(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}}{k_\rho}\right\}$
G_{yy}	$-\sin^2\psi\,\mathcal{S}_0\left\{V_i^{\mathrm{TM}}\right\} - \cos^2\psi\,\mathcal{S}_0\left\{V_i^{\mathrm{TE}}\right\} - \frac{\cos(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{V_i^{\mathrm{TM}} - V_i^{\mathrm{TE}}}{k_\rho}\right\}$
G_{yz}	$\frac{\eta_0}{jk_0\varepsilon(z')}\sin\psi\,\mathcal{S}_1\left\{k_\rho V_v^{\mathrm{TM}}\right\}$
G_{zx}	$\frac{\eta_0}{jk_0\varepsilon(z)}\cos\psi\,\mathcal{S}_1\left\{k_\rho I_i^{\mathrm{TM}}\right\}$
G_{zy}	$\frac{\eta_0}{jk_0\varepsilon(z)}\sin\psi\,\mathcal{S}_1\left\{k_\rho I_i^{\mathrm{TM}}\right\}$
G_{zz}	$-\frac{\eta_0^2}{jk_0^2\varepsilon(z)\varepsilon(z')}\mathcal{S}_0\left\{k_\rho^2 I_v^{\mathrm{TM}}\right\} - \frac{\eta_0}{jk_0\varepsilon(z)}\delta(\boldsymbol{\rho})\delta(z-z')$

Table 3.2: Scalar Green functions for computation of electric field due to an electric current source. Adapted from [2]

GF	$\kappa = \mathrm{HJ}$
G_{xx}	$-\frac{\sin(2\psi)}{2}\mathcal{S}_0\left\{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}\right\} + \frac{\sin(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}}{k_\rho}\right\}$
G_{xy}	$\cos^2\psi\,\mathcal{S}_0\left\{I_i^{\mathrm{TE}}\right\} + \sin^2\psi\,\mathcal{S}_0\left\{I_i^{\mathrm{TM}}\right\} - \frac{\cos(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}}{k_\rho}\right\}$
G_{xz}	$-\frac{\eta_0}{jk_0\varepsilon(z)}\sin\psi\,\mathcal{S}_1\left\{k_\rho I_v^{\mathrm{TM}}\right\}$
G_{yx}	$-\sin^2\psi\,\mathcal{S}_0\left\{I_i^{\mathrm{TE}}\right\} - \cos^2\psi\,\mathcal{S}_0\left\{I_i^{\mathrm{TM}}\right\} - \frac{\cos(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}}{k_\rho}\right\}$
G_{yy}	$\frac{\sin(2\psi)}{2}\mathcal{S}_0\left\{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}\right\} - \frac{\sin(2\psi)}{\rho}\mathcal{S}_1\left\{\frac{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}}{k_\rho}\right\}$
G_{yz}	$\frac{\eta_0}{jk_0\varepsilon(z')}\sin\psi\,\mathcal{S}_1\left\{k_\rho I_v^{\mathrm{TM}}\right\}$
G_{zx}	$\frac{1}{jk_0\eta_0\mu}\sin\psi\,\mathcal{S}_1\left\{k_\rho V_i^{\mathrm{TE}}\right\}$
G_{zy}	$-\frac{1}{jk_0\eta_0\mu}\cos\psi\,\mathcal{S}_1\left\{k_\rho V_i^{\mathrm{TE}}\right\}$
G_{zz}	.

Table 3.3: Scalar Green functions for computation of magnetic field due to an electric current source. Adapted from [2]

3.1.5 Introducing Potential Functions

As evident in Tables 3.2 and 3.3, the spatial domain DGFs that directly describe fields are tedious. An alternative formulation using vector potential functions considerably reduces the computation complexity. The approach is commonly known as mixed potentials integral equation (MPIE) formulation. From

the Maxwell's equations, we know that,

$$\nabla \cdot \mu \mathbf{H} = 0, \tag{3.40}$$

and using the vector identity, $\nabla \cdot (\nabla \times \mathbf{A}) \equiv 0$, we define the magnetic vector potential, A such that:

$$\mathbf{H} = \frac{1}{\mu} \nabla \times \mathbf{A}. \tag{3.41}$$

Through (3.2a) and (3.41), we write:

$$\nabla \times (\mathbf{E} + j\omega \mathbf{A}) = 0, \tag{3.42}$$

and knowing the vector identity, $\nabla \times (-\nabla \phi) \equiv 0$, the electric field is expressed as:

$$\mathbf{E} = j\omega \mathbf{A} - \nabla \phi. \tag{3.43}$$

The scalar potential function ϕ can be related to A by the Lorenz gauge, which states that:

$$\nabla \cdot \mathbf{A} = -j\omega \mu \varepsilon \phi. \tag{3.44}$$

The electric field in (3.43) is then expressed only as a function of $\mathbf{A}$,

$$\mathbf{E} = \frac{j\omega}{k^2} \left(k^2 + \nabla \nabla \cdot \right) \mathbf{A}. \tag{3.45}$$

For a current distribution $\mathbf{J}$, the vector potential $\mathbf{A}$ can also be described by a DGF in a similar fashion to (3.1):

$$\mathbf{A} = \int_{r'} \underline{\underline{\mathbf{G}}}^{\mathrm{A}}(\mathbf{r}|\mathbf{r}') \mathbf{J}(\mathbf{r}') \, \mathrm{d}r', \tag{3.46}$$

where $\underline{\underline{\mathbf{G}}}^{\mathrm{A}}$ is the vector potential DGF. Following the formulation of [8], the spectral domain form is,

$$\underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{A}} = (\hat{\mathbf{u}}\hat{\mathbf{u}} + \hat{\mathbf{v}}\hat{\mathbf{v}})\,\tilde{G}^{\mathrm{A}}_{vv} + \hat{\mathbf{z}}\hat{\mathbf{u}}\tilde{G}^{\mathrm{A}}_{zu} + \hat{\mathbf{z}}\hat{\mathbf{z}}\tilde{G}^{\mathrm{A}}_{zz}. \tag{3.47}$$

The vector potential DGF is related to its fields counterpart through the magnetic field in (3.41) and written as:

$$\underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{HJ}} = \frac{1}{\mu}\nabla \times \underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{A}} \tag{3.48}$$

where isotropic permeability is assumed. To find the components of $\underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{A}}$, we first note that the ∇ operator used in the curl operation, which was initially defined in rectangular coordinates, is re-written in the spectral domain form as $\tilde{\nabla} = -jk_\rho\hat{\mathbf{u}} + \hat{\mathbf{z}}d/dz$. Therefore, (3.48) becomes,

$$\begin{aligned}
\nabla \times \underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{A}} =& \left(-jk_\rho\hat{\mathbf{u}} + \hat{\mathbf{z}}d/dz\right) \times \left[(\hat{\mathbf{u}}\hat{\mathbf{u}} + \hat{\mathbf{v}}\hat{\mathbf{v}})\,\tilde{G}^{\mathrm{A}}_{vv} + \hat{\mathbf{z}}\hat{\mathbf{u}}\tilde{G}^{\mathrm{A}}_{zu} + \hat{\mathbf{z}}\hat{\mathbf{z}}\tilde{G}^{\mathrm{A}}_{zz}\right] \\
=& -jk_\rho\tilde{G}^{\mathrm{A}}_{vv}\,(\hat{\mathbf{u}}\times\hat{\mathbf{v}})\hat{\mathbf{v}} + \frac{\mathrm{d}\tilde{G}^{\mathrm{A}}_{vv}}{\mathrm{d}z}\,(\hat{\mathbf{z}}\times\hat{\mathbf{u}})\hat{\mathbf{u}} - jk_\rho\tilde{G}^{\mathrm{A}}_{zu}\,(\hat{\mathbf{u}}\times\hat{\mathbf{z}})\hat{\mathbf{u}} \\
& - jk_\rho\tilde{G}^{\mathrm{A}}_{zz}\,(\hat{\mathbf{u}}\times\hat{\mathbf{z}})\hat{\mathbf{z}} \\
=& -jk_\rho\tilde{G}^{\mathrm{A}}_{vv}\,\hat{\mathbf{z}}\hat{\mathbf{v}} + \frac{\mathrm{d}\tilde{G}^{\mathrm{A}}_{vv}}{\mathrm{d}z}\,\hat{\mathbf{v}}\hat{\mathbf{u}} + jk_\rho\tilde{G}^{\mathrm{A}}_{zu}\,\hat{\mathbf{v}}\hat{\mathbf{u}} + jk_\rho\tilde{G}^{\mathrm{A}}_{zz}\,\hat{\mathbf{v}}\hat{\mathbf{z}}.
\end{aligned} \tag{3.49}$$

Using (3.48) and comparing the dyads of (3.26b) and (3.49), we obtain the scalar components of the vector DGF, $\underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{A}}$:

$$j\omega\tilde{G}^{\mathrm{A}}_{vv} = V_i^{\mathrm{TE}}, \tag{3.50a}$$

$$j\omega\mu\varepsilon(z')\tilde{G}^{\mathrm{A}}_{zz} = I_v^{\mathrm{TM}}, \tag{3.50b}$$

$$\frac{\mathrm{d}\tilde{G}^{\mathrm{A}}_{vv}}{\mathrm{d}z} + jk_\rho\tilde{G}^{\mathrm{A}}_{zu} = -\mu I_v^{\mathrm{TM}}. \tag{3.50c}$$

Substituting (3.50a) into (3.50c) and invoking the telegrapher's equation, (3.20a), we obtain:

$$\begin{aligned} \frac{1}{j\omega}\left[-jk_z^{\mathrm{TE}} Z^{\mathrm{TE}} I_i^{\mathrm{TE}}\right] + jk_\rho \tilde{G}_{zu}^{\mathrm{A}} &= -\mu I_v^{\mathrm{TM}} \\ \tilde{G}_{zu}^{\mathrm{A}} &= \frac{\mu}{jk_\rho}\left(I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}\right), \end{aligned} \tag{3.51}$$

where $Z^{\mathrm{TE}} = \omega\mu/k_z^{\mathrm{TE}}$.

3.1.6 Spatial Domain DGFs

The spectral domain vector potentials defined in (3.47) can be conveniently written in a matrix form containing scalar elements,

$$\underline{\underline{\tilde{\mathbf{G}}}}^{\mathrm{A}} = \begin{bmatrix} \tilde{G}_{vv}^{\mathrm{A}} & & \\ & \tilde{G}_{vv}^{\mathrm{A}} & \\ \frac{k_x}{k_\rho}\tilde{G}_{zu}^{\mathrm{A}} & \frac{k_y}{k_\rho}\tilde{G}_{zu}^{\mathrm{A}} & \tilde{G}_{zz}^{\mathrm{A}} \end{bmatrix}. \tag{3.52}$$

A Fourier inversion following a similar procedure used to convert fields DGFs (3.35)-(3.39) yields the spatial domain potential DGFs expressed in terms of TL-GFs, which are listed in Table 3.4.

The above formulation is applicable to only multilayer structures in which the layers are stacked along the x-y plane. The formulation allows insertion of arbitrarily oriented electric current source that can be placed anywhere in the structure. Once the magnetic vector potential, **A**, is computed using (3.46), the corresponding electric and magnetic fields can be found through (3.45) and (3.41) respectively.

$G^{\mathrm{A}}_{xx} = G^{\mathrm{A}}_{yy}$	$\dfrac{1}{j\omega}\mathcal{S}_0\left\{V_i^{\mathrm{TE}}\right\}$
G^{A}_{zx}	$j\mu\cos\phi\,\mathcal{S}_1\left\{\dfrac{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}}{k_\rho}\right\}$
G^{A}_{zy}	$j\mu\sin\phi\,\mathcal{S}_1\left\{\dfrac{I_i^{\mathrm{TE}} - I_i^{\mathrm{TM}}}{k_\rho}\right\}$
G^{A}_{zz}	$\dfrac{1}{j\omega\mu\varepsilon(z')}\mathcal{S}_0\left\{I_v^{\mathrm{TM}}\right\}$

Table 3.4: Scalar potential Green functions expressed in terms of TLGFs

3.2 Computation of Transmission Line Green Functions

The voltage and current at any point in a transmission line can be obtained [78] using simple network analysis techniques. Each layer in the multilayer structure is represented by a section in the transmission line which is specified by its characteristic impedance Z and the propagation constant k_z. As an example, an equivalent TL network for a semiconductor heterostructure that forms the substrate of the modern transistors is shown in Fig. 3.3. The layers numbered 2 and 3 in Fig. 3.3a are commonly made up of group III-V materials such as gallium nitride (GaN) and their subsequent alloys such aluminum gallium nitride (AlGaN). Based on the device type, the top layer labeled 1, can represent either freespace or a perfectly conducting material. The thin layer which is sandwiched between the semiconductor layers is known as a two-dimensional electron gas (2DEG) which can be visualized as a highly conductive sheet of electrons that is known to exhibit extraordinary electromagnetic properties such as highly confined surface waves [83, 84]. In this section, we formulate the TLGFs for various configuration of the 2DEG and compute the respective potential Green functions.

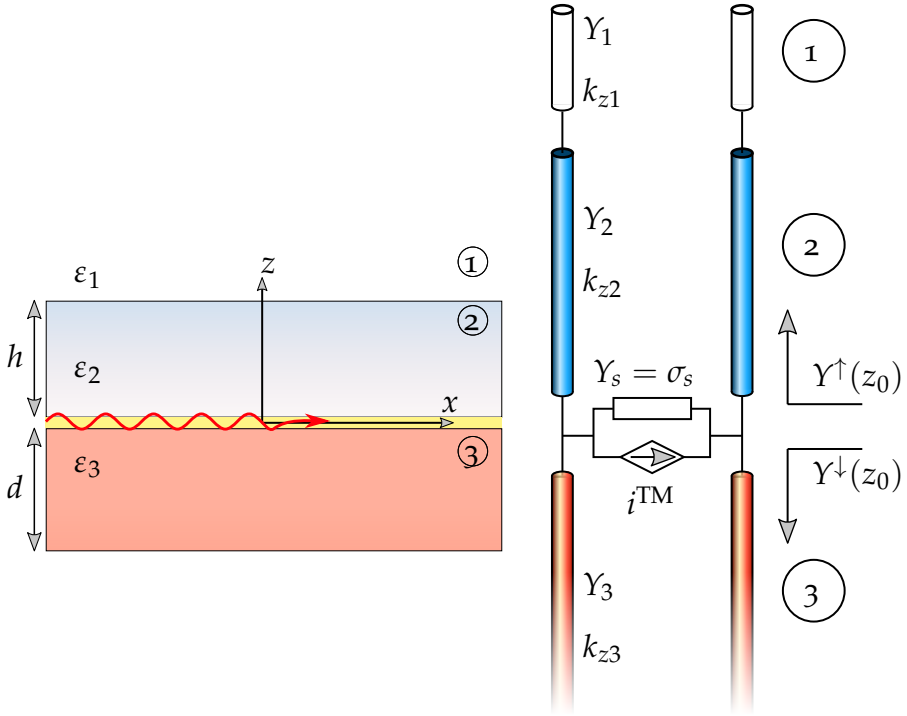


Figure 3.3: (a) Multilayer structure typically found in a high electron mobility transistor, (b) equivalent transmission line network

3.2.1 Lossy Half-space

We start with a case in which the 2DEG occupies the whole lower half-space. Such a case is an example of the classical Sommerfeld half-space problem that has been actively researched in the past century [85, 86]. An electric dipole of arbitrary orientation is placed in air at a height z' from the interface where the origin is also centered. The observation point, at first also lies in air. To find the component G_{xx}^{A}, we require TE-mode voltage of an equivalent, current excited TL network which is expressed using (3.25a) and (3.25b) [78, Sec. 2.4] [8],

$$V_i^{\mathrm{TE}}(z,z') = \frac{Z_1^{\mathrm{TE}}}{2}\left[\mathrm{e}^{-jk_{z1}|z-z'|} + \Gamma^{\downarrow,\mathrm{TE}}\,\mathrm{e}^{-jk_{z1}(z+z')}\right], \tag{3.53}$$

where $\Gamma^{\downarrow,\mathrm{TE}} = (Z^{\mathrm{TE}}_{\mathrm{2DEG}} - Z^{\mathrm{TE}}_1)/(Z^{\mathrm{TE}}_{\mathrm{2DEG}} + Z^{\mathrm{TE}}_1)$, $Z^{\mathrm{TE}}_{\mathrm{2DEG}}$ and Z^{TE}_1 are the TE mode impedances of the 2DEG and freespace regions respectively, given by (3.21b). The scalar potential GF, G^{A}_{xx} is then expressed as:

$$G^{\mathrm{A}}_{xx} = \mathcal{S}_0 \left\{ \frac{\mathrm{e}^{-jk_{z1}|z-z'|}}{2jk_{z1}} \right\} + \mathcal{S}_0 \left\{ \frac{\Gamma^{\downarrow,\mathrm{TE}}\, \mathrm{e}^{-jk_{z1}(z+z')}}{2jk_{z1}} \right\}. \tag{3.54}$$

Following similar steps, G^{A}_{zz} is found using the TM-mode current of a voltage excited network. For the components G^{A}_{zx} and G^{A}_{zy}, both the TM and TE-mode currents need to be considered for a current excited network. The final expressions are:

$$G^{\mathrm{A}}_{zz} = \mathcal{S}_0 \left\{ \frac{\mathrm{e}^{-jk_{z1}|z-z'|}}{2jk_{z1}} \right\} - \mathcal{S}_0 \left\{ \frac{\Gamma^{\downarrow,\mathrm{TM}}\, \mathrm{e}^{-jk_{z1}(z+z')}}{2jk_{z1}} \right\} \tag{3.55}$$

and,

$$G^{\mathrm{A}}_{zx} = \frac{j\mu}{2} \cos\phi\, \mathcal{S}_1 \left\{ \frac{\Gamma^{\downarrow,\mathrm{TM}} - \Gamma^{\downarrow,\mathrm{TE}}}{k_\rho} \mathrm{e}^{-jk_{z1}(z+z')} \right\}. \tag{3.56}$$

The first term in (3.54) and (3.55) is the Sommerfeld identity which can be written as the freespace Green function of a point source,

$$4\pi\mathcal{S}_0 \left\{ \frac{\mathrm{e}^{-jk_{z1}|z-z'|}}{2jk_{z1}} \right\} = \frac{\mathrm{e}^{-jk_1 R}}{4\pi R}. \tag{3.57}$$

The plasma frequency ω_p of the 2DEG lies in the terahertz frequency at which the dielectric function vanishes to zero. At frequencies lower than ω_p, the real part is negative and the imaginary part, which expresses loss is highly dependent on temperature. Here, we assume a dielectric constant of $-8.114 - j8.159$ for a GaN/AlGaN heterostructure at 5.600 THz, and observed at room temperature (295 K).

For a horizontally oriented electric dipole, only the spatial domain GFs in G^{A}_{xx}

(3.54), G^{A}_{zx} (3.56) exist, and are plotted in Fig. 3.4 by excluding the freespace GF part. To compute the SIs, real-axis integration using a combination of *tanh-sinh*

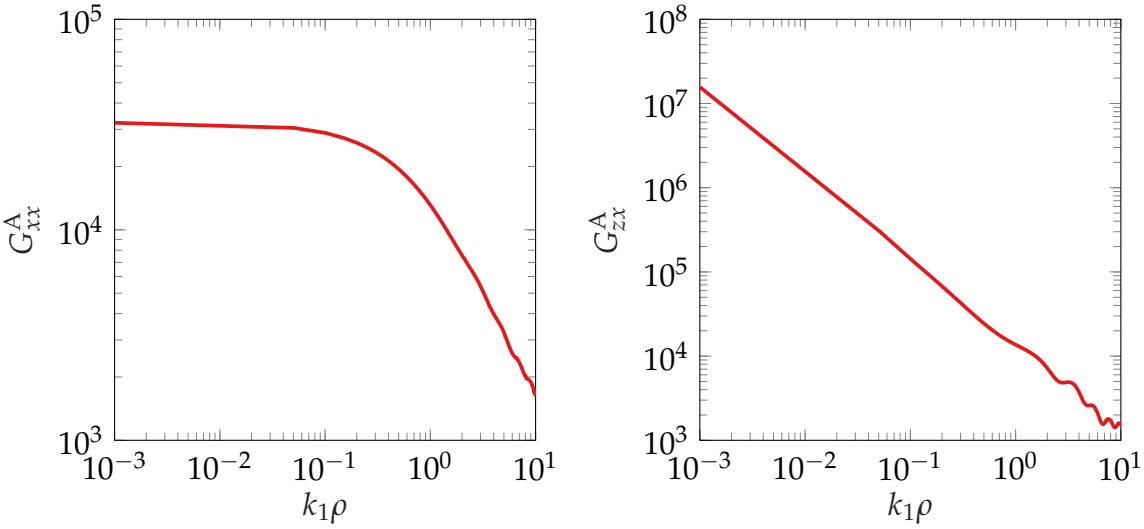


Figure 3.4: Components of potential DGF for a 2DEG half-space (a) G^{A}_{xx} (3.54), (b) G^{A}_{zx} (3.56)

quadrature [87, 88, 89] and partition extrapolation [90]. Algorithmic pseudo-code for both techniques is discussed in [91], following which the integrals are computed here.

3.2.2 Thin Sheet

Next we consider the case of an infinitesimally thin sheet of 2DEG located on the x-y plane and surrounded by freespace illustrated along with its TL model in Fig. 3.5. A horizontal electric dipole is placed at a height h above the sheet. The surface conductivity of the thin sheet is expressed in terms of a Drude-type

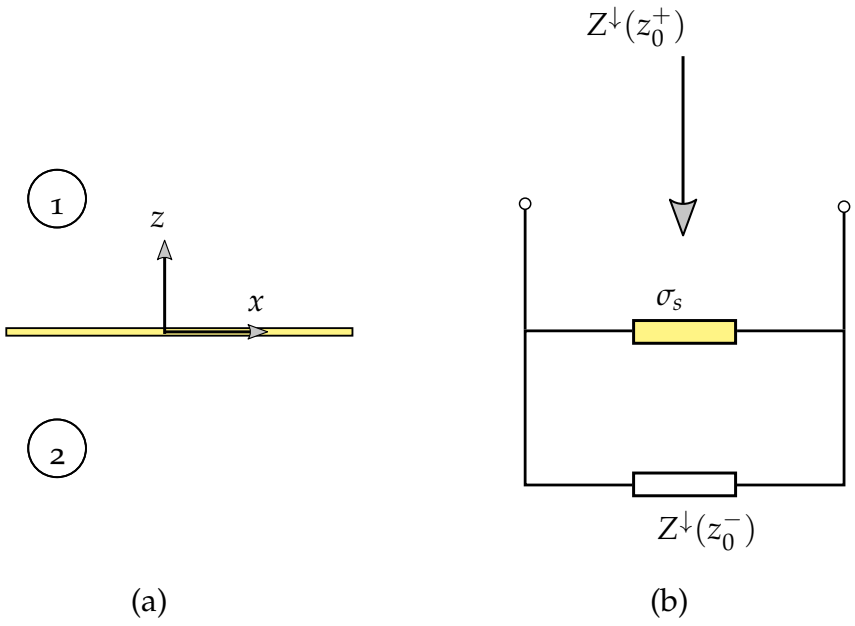


Figure 3.5: (a) 2DEG sheet in surrounded by freespace, (b) equivalent transmission line circuit consisting of a shunt impedance

model:

$$\sigma_s = \frac{N_s e^2 \tau}{m^*} \frac{1}{1 + j\omega\tau}, \tag{3.58}$$

where N_s is the free electron concentration at the interface, e is the electron charge, m^* is the effective electron mass in the heterostructure, τ is the scattering time of electrons. The dielectric function is subsequently expressed as,

$$\varepsilon(\omega) = 1 - j\frac{\sigma_s}{\omega\Delta\varepsilon_r\varepsilon_0}, \tag{3.59}$$

in which Δ is the thickness of the 2DEG region, ε_r is the dielectric constant of the GaN layer which acts as a substrate and ε_0 is the freespace permittivity. For a 5 nm thick GaN/AlGaN based 2DEG, where $N_s = 7.5 \times 10^{12}\,\mathrm{cm}^{-2}$ and $\varepsilon_r = 9.6$ [92], the surface conductivity and dielectric function are plotted in Fig. 3.6 at

room temperature. The equivalent transmission line network used to calculate

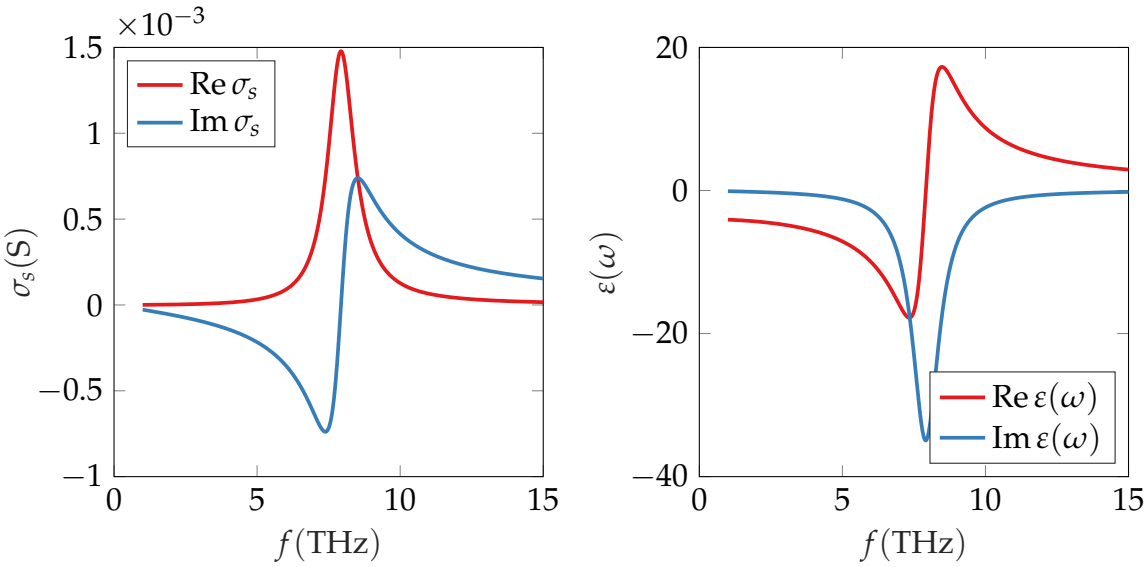


Figure 3.6: Electrical properties of a GaN/AlGaN 2DEG at 295 K, (a) surface conductivity (3.58) (b) dielectric function (3.59)

the reflection coefficients consists of a simple circuit of two parallel impedances [93, 94, 95]. Assuming the source lies on the top of the 2DEG, the down-looking impedance of the shunt configuration circuit is,

$$Z^{\downarrow}(z_0^+) = \frac{Z^{\downarrow}(z_0^-)}{1 + \sigma_s \, Z^{\downarrow}(z_0^-)} \tag{3.60}$$

where it is noted that σ_s has the dimensions of admittance, and $Z^{\downarrow}(z_0^-)$ is the down-looking effective impedance as seen from the bottom of the sheet. In the case considered, $Z^{\downarrow}(z_0^-) = Z_0$ where Z_0 is the freespace intrinsic impedance. The reflections needed in (3.54)-(3.56) are listed in Table 3.5.

$\Gamma^{\downarrow,\mathrm{TE}}$	$\Gamma^{\downarrow,\mathrm{TM}}$
$\dfrac{k_{z1}-\omega\mu_1\sigma_s}{k_{z1}+\omega\mu_1\sigma_s}$	$\dfrac{\omega\varepsilon_1-\sigma_s k_{z1}}{\omega\varepsilon_1+\sigma_s k_{z1}}$

Table 3.5: Reflection coefficients for 2DEG sheet in freespace

3.2.2.1 Location of Poles

The integrands of the SIs in (3.54)-(3.56) contain poles that are associated with the zeros of the integrand denominator. For a thin sheet in free-space, the poles can be analytically found by applying using the transverse resonance (TR) condition [78, Sec. 2.4], which requires that the total impedance at any point ($z = 0$ in this case) must be zero. Using the TM parameters given in Table 3.5 and (3.60), we write:

$$Z^{\downarrow} + Z^{\uparrow} = 0, \tag{3.61a}$$

$$Z_1 + Z_1(Z_1\sigma_s + 1) = 0. \tag{3.61b}$$

The expression is obtained by substituting (3.21a) into (3.61b) [89],

$$k_{\mathrm{P}}^{\mathrm{TM}} = k_1\sqrt{1-\left(\frac{2}{\eta_0\sigma_s}\right)^2} \tag{3.62}$$

Similarly the TE-mode pole is expressed as,

$$k_{\mathrm{P}}^{\mathrm{TE}} = k_1\sqrt{1-\left(\frac{\eta_0\sigma_s}{2}\right)^2} \tag{3.63}$$

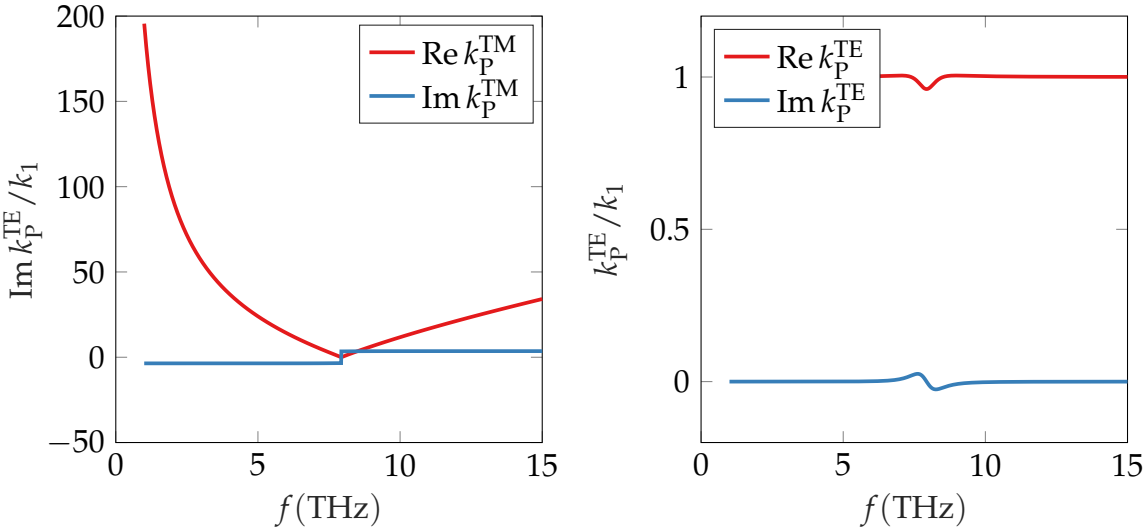
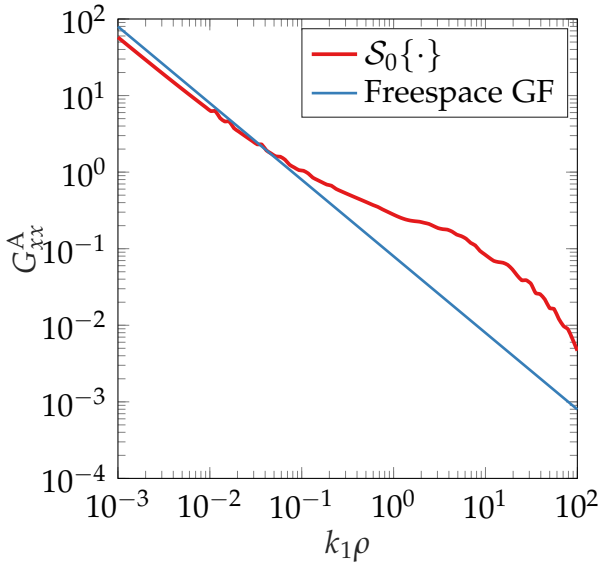
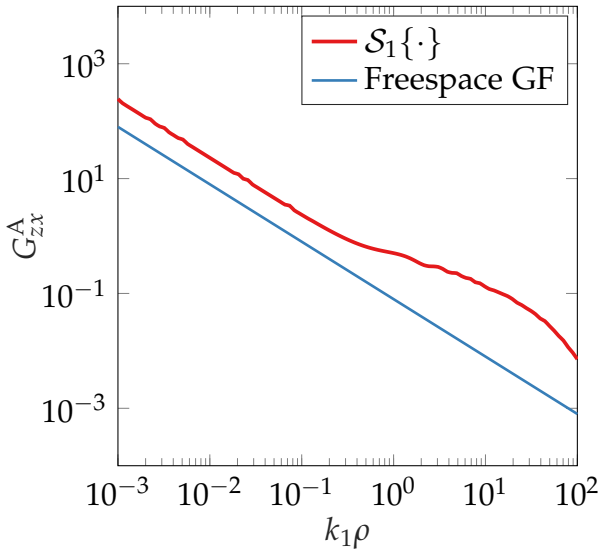


Figure 3.7: Distribution of analytically calculated poles in the complex plane (a) TM mode (3.62) (b) TE mode (3.63)

A mode is termed as proper only when the pole lies in the fourth quadrant of the complex plane. As evident from Figs. 3.6 and 3.7a, proper TM modes only exist when the real part of dielectric function of the 2DEG is negative which also corresponds to a negative imaginary part of the conductivity. On the other hand, a proper TE mode exists when the real part of dielectric function is positive which is shown in Fig. 3.7b. Figures 3.8 and 3.9 show the components of the potential GF, G^{A}, due to a horizontally oriented electric dipole located on the top side of the sheet. The observation point also lies on the same side as the source. The freespace GF is also plotted in each figure as a reference. In the near-field region that corresponds to normalized lateral distances less than 1×10^{-1}, the overall field contribution on the sheet mainly comes from the source since the slope of the curve is the same as in freespace GF. The change in slope in the intermediate region shows the existence of plasmon like surface waves.

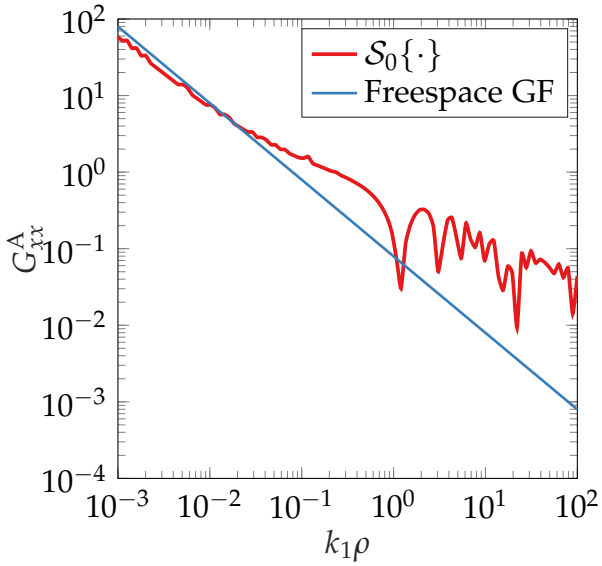


(a)

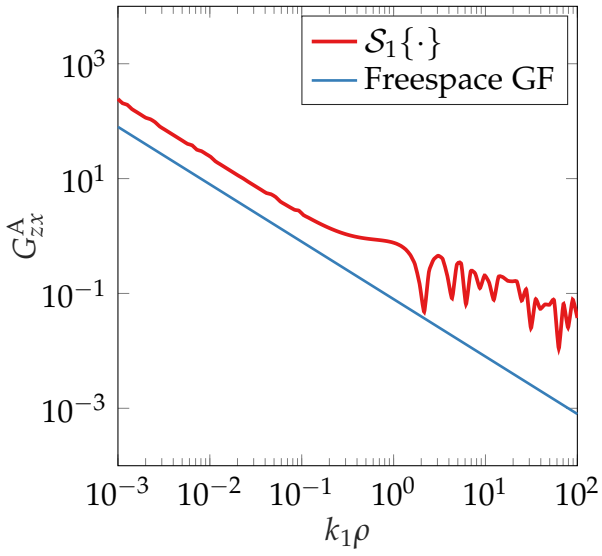


(b)

Figure 3.8: Components of G^{A} computed for a GaN/AlGaN based 2DEG sheet suspended in freespace. The surface conductivity of the sheet is $\sigma_s = 7.6 \times 10^{-5} - \mathrm{j}2.98 \times 10^{-3}$ S corresponding to 5.6 THz at room temperature (295 K). (a) G^{A}_{xx}, (b) G^{A}_{zx}



(a)



(b)

Figure 3.9: Components of G^{A} computed for a GaN/AlGaN based 2DEG sheet suspended in freespace. The cryogenic surface conductivity of the sheet is $\sigma_s = 7.6 \times 10^{-8} - \mathrm{j}2.98 \times 10^{-3}\,\mathrm{S}$ corresponding to 5.6 THz at a temperature (3 K). (a) G^{A}_{xx}, (b) G^{A}_{zx}

The propagation length of the surface waves fundamentally depends on the loss which manifests in the form of real part of the conductivity. In Fig. 3.8 in which the GFs are computed at room temperature, the plasmonic contribution decays substantially faster owing to the lossy nature of the 2DEG at room temperature. On the other hand, when the 2DEG is kept in a cryogenic environment, the real part is decreased by at least three order of magnitude resulting in an extremely low-loss setup that permits much larger propagation lengths. In Fig. 3.9, the GFs are computed at cryogenic temperature and the increase in propagation length is noticeable in the intermediate to far field region.

In this section, a transmission line network based technique is discussed through which spectral and subsequently spatial domain potential GFs have been derived for a multilayer structure, with the aim to explore surface wave propagation in the terahertz frequency range. Of particular interest is an infinitesimally thin sheet of free-charges known as a 2DEG, for which components of the potential GFs are computed. Through the location of the analytically poles in the complex plane, it was shown that when the surface conductivity of the sheet has a negative imaginary part, proper TM modes exist. The required condition holds true for III-V heterostructures in the terahertz frequency range. Furthermore, the GFs exhibit plasmonic behavior in the intermediate to far-field region beyond the plasmons decay due to the loss present in the sheet.

4. AN INTEGRAL EQUATION SCHEME FOR THIN PLANAR SHEETS

An integral equation formulation for a thin dielectric sheet is presented that is derived using the surface equivalence theorem. Unlike other schemes that involve some sort of approximation to compute fields for thin objects, our method provides an exact solution when the thickness is infinitesimal. Numerical results are presented to illustrate the scattering properties of a thin sheet and compared with other schemes.

4.1 Background

The emergence of high-precision nanoscale fabrication techniques has led to an increased interest in two-dimensional (2D) materials and electronic systems of late, especially in the terahertz frequency regime. One particular intriguing example is a two-dimensional electron gas (2DEG) found in the multilayer stack of semiconductor structures like high-electron mobility transistors (HEMTs), that exhibits remarkable electrical properties such as high electron mobility along with very high with free electron densities. The 2DEG is extremely thin as compared to other layer thicknesses in the stack and therefore, its scattering properties can be found by modeling it as a two-dimensional plasma. An interaction between an external electromagnetic radiation and plasma results in 2D plasmons (surface waves). Historically, wave analysis of thin objects that are not perfectly conductors has involved approximation methods such as the Leontovich boundary condition [96, 97] that relates the tangential electric and magnetic fields on the object surface:

$$\mathbf{E}_{tan} = Z\,\hat{\mathbf{n}} \times \mathbf{H}, \tag{4.1}$$

where Z is the surface impedance and $\hat{\mathbf{n}}$ is the outward unit vector, normal to the surface. The accuracy of the boundary condition in (4.1) is determined by the complexity and material of the object.

Here, we formulate the scattering response of an infinitesimally thin flat layer of plasma surrounded by free-space using the surface equivalence theorem.

4.2 Theory

4.2.1 Surface Equivalence Theorem

The field computation due to sources present in an inhomogeneous environment can be simplified using the surface equivalence theorem in which the actual sources are replaced by a set of fictitious, yet equivalent sources illustrated in Fig. 4.1. The solution can be divided into two homogeneous spaces, one internal and the other exterior. Considering the configuration of Fig. 4.1a, the external equivalent is a homogeneous space composed of the external material. The total fields in the external region due to equivalent electric and magnetic surface currents, $\mathbf{J}_s$ and $\mathbf{M}_s$ respectively, for the case when the object is collapsed to a flat sheet can be expressed as:

$$\begin{aligned}\mathbf{E}_1 &= \mathbf{E}^i + \mathbf{E}_1^{scat} \\ &= -\frac{\omega}{4k_1^2}\left(k_1^2 + \nabla\nabla\cdot\right)\int_C \mathbf{J}_s(\boldsymbol{\rho}')H_0^{(2)}(k_1|\boldsymbol{\rho}-\boldsymbol{\rho}'|)\,\mathrm{d}l' \\ &\quad - \frac{1}{4\varepsilon j}\nabla\times\int_l \mathbf{M}_s(\boldsymbol{\rho}')H_0^{(2)}(k_1|\boldsymbol{\rho}-\boldsymbol{\rho}'|)\,\mathrm{d}l' + \mathbf{E}^i\end{aligned} \tag{4.2}$$

$$
\begin{aligned}
\mathbf{H}_1 &= \mathbf{H}^i + \mathbf{H}_1^{scat} \\
&= \frac{1}{4j}\nabla \times \int_l \mathbf{J}_s(\boldsymbol{\rho}') H_0^{(2)}(k_1|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\,\mathrm{d}l' \\
&- \frac{\omega}{4k_1^2}\left(k_1^2 + \nabla\nabla\cdot\right) \int_l \mathbf{M}_s(\boldsymbol{\rho}') H_0^{(2)}(k_1|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\,\mathrm{d}l' + \mathbf{H}^i
\end{aligned}
\tag{4.3}
$$

where $\mathbf{E}^i$ and $\mathbf{H}^i$ are the incident electric and magnetic fields that are known, k_1 is the free-space propagation constant, $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ are the position vectors of any point from the origin and source respectively, and $H_0^2(\cdot)$ is the zero order Hankel function of the second kind. Similarly, fields for the interior region that only contain the scattered part are,

$$
\begin{aligned}
\mathbf{E}_2 &= \mathbf{E}_2^{scat} \\
&= -\frac{\omega}{4k_2^2}\left(k_2^2 + \nabla\nabla\cdot\right) \int_C \left(-\mathbf{J}_s(\boldsymbol{\rho}')\right) H_0^{(2)}(k_2|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\,\mathrm{d}l' \\
&- \frac{1}{4j}\nabla \times \int_l \left(-\mathbf{M}_s(\boldsymbol{\rho}')\right) H_0^{(2)}(k_2|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\,\mathrm{d}l'
\end{aligned}
\tag{4.4}
$$

$$
\begin{aligned}
\mathbf{H}_2 &= \mathbf{H}_1^{scat} \\
&= \frac{1}{4j}\nabla \times \int_l \left(-\mathbf{J}_s(\boldsymbol{\rho}')\right) H_0^{(2)}(k_2|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\,\mathrm{d}l' \\
&- \frac{\omega}{4k_2^2}\left(k_2^2 + \nabla\nabla\cdot\right) \int_l \left(-\mathbf{M}_s(\boldsymbol{\rho}')\right) H_0^{(2)}(k_2|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\,\mathrm{d}l'
\end{aligned}
\tag{4.5}
$$

where k_2 is the propagation constant in the interior region.

4.2.2 Surface Integral Equation

4.2.2.1 TM_z Polarization

Next we consider a TM_z excited planar dielectric sheet lying along the x-axis and apply the surface equivalence theorem to find the electric and magnetic

currents on the sheet. A plane wave propagating along the direction $\mathbf{k}$ with electric field $\mathbf{E}$ polarized along the z direction as shown in Fig. 4.2a is incident on the dielectric surface at an angle ϕ_i.

$$\mathbf{J}_s = \hat{\mathbf{n}} \times \mathbf{H} = \hat{\mathbf{z}}\, \mathrm{J}(\zeta), \tag{4.6a}$$

$$\mathbf{M}_s = -\hat{\mathbf{n}} \times \mathbf{E} = \hat{\mathbf{x}}\, \mathrm{M}(\zeta) \tag{4.6b}$$

where the normal unit vector $\hat{\mathbf{n}}$ is in the y direction and ζ depends on x and y coordinates. To find the surface currents, we set up an homogeneous equivalent problem first for the region outside the dielectric sheet due to an incident field,

$$\mathbf{E}^i = \hat{\mathbf{z}}\; E_0 \mathrm{e}^{-jk_1(x\cos\phi_i - y\sin\phi_i)} \tag{4.7}$$

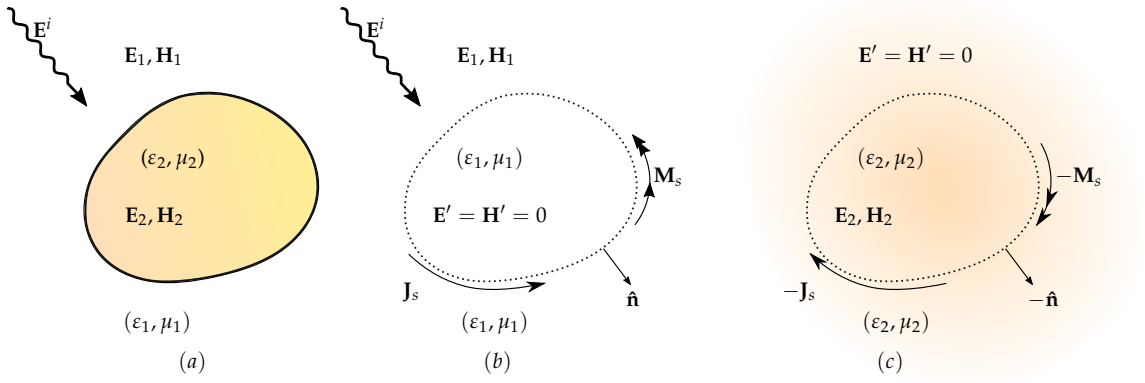
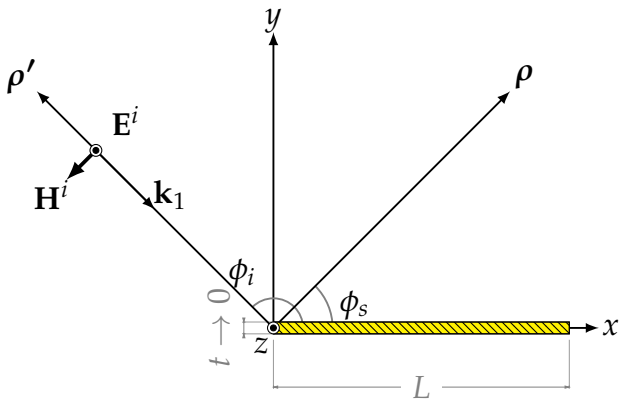


Figure 4.1: (a) Actual and its equivalent model for the (b) external and, (c) internal region

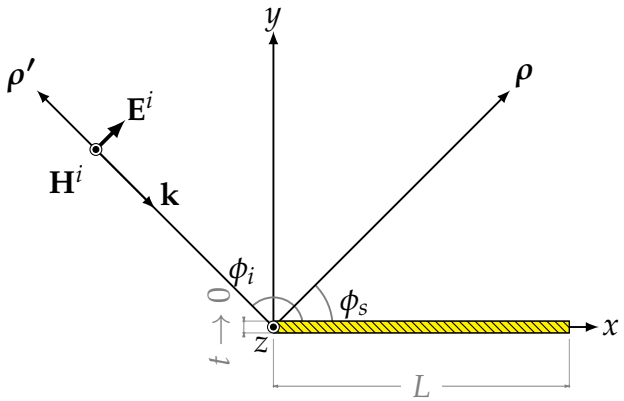
with E_0 the amplitude of the incoming plane wave. We now express the scattered fields in terms of potentials as

$$\mathbf{E}_1^{scat} = -\frac{j\omega}{k_1^2}\left(k_1^2 + \nabla\nabla\cdot\right)\mathbf{A}, \tag{4.8a}$$

$$\mathbf{H}_1^{scat} = -\frac{j\omega}{k_1^2}\left(k_1^2 + \nabla\nabla\cdot\right)\mathbf{F}, \tag{4.8b}$$



(a)



(b)

Figure 4.2: Thin plasma sheet with under (a) TM_z and (b) TE_z polarizations

where $\mathbf{A}$ and $\mathbf{F}$ are the magnetic and electric vector potentials respectively, given by:

$$\mathbf{A} = \frac{\mu}{4j} \int_l \mathbf{J}_s(\boldsymbol{\rho}') H_0^{(2)}(k_1|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\, \mathrm{d}l', \tag{4.9a}$$

$$\mathbf{F} = \frac{\varepsilon}{4j} \int_l \mathbf{M}_s(\boldsymbol{\rho}') H_0^{(2)}(k_1|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\, \mathrm{d}l', \tag{4.9b}$$

The scattered electric field off a flat plate oriented along the x-axis can be simply written in a scalar form due to a z-directed incident wave,

$$\begin{aligned} E_1^{scat} &= -j\omega \mathrm{A}_z \\ &= -\frac{\omega\mu}{4j} \int_l J_z(x') H_0^{(2)}(k_1|\boldsymbol{\rho} - \boldsymbol{\rho}'|)\, \mathrm{d}l' \end{aligned} \tag{4.10}$$

Likewise, the magnetic field that is tangential to the plate is,

$$\begin{aligned} H_{1,x}^{scat} &= -\frac{j\omega}{k_1^2} \left(k_0^2 + \frac{\partial^2}{\partial x^2} \right) \mathrm{F}_x \\ &= -\frac{j\omega}{k_1^2} \left(k_1^2 + \frac{\partial^2}{\partial x^2} \right) \int_l M_x(x') H_0^{(2)}(k_1|x - x'|)\, \mathrm{d}l'. \end{aligned} \tag{4.11}$$

Using a similar procedure, we set up an interior equivalent with the currents reversing the signs. The total fields for the interior region only contain the scattered fields.

$$E_2^{scat} = -\frac{\omega\mu}{4j} \int_l -J_z(x') H_0^{(2)}(k_2|x - x'|)\, \mathrm{d}l' \tag{4.12a}$$

$$H_{2,x}^{scat} = -\frac{j\omega}{k_2^2} \left(k_2^2 + \frac{\partial^2}{\partial x^2} \right) \int_l -M_x(x') H_0^{(2)}(k_2|x - x'|)\, \mathrm{d}l' \tag{4.12b}$$

In order to find the electric and magnetic currents, we apply the boundary conditions at the interface ensuring the continuity of tangential component of the fields. At the interface:

$$\hat{\mathbf{n}} \times (\mathbf{E}_1 - \mathbf{E}_2) = \mathbf{0} \tag{4.13a}$$

$$\hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{0} \tag{4.13b}$$

Using (4.13a), (4.10) and (4.12b) we obtain:

$$E_i = \frac{\omega\mu}{4} \int_c J_z(x') \left[H_0^{(2)}(k_1|x - x'|) + H_0^{(2)}(k_2|x - x'|) \right] \mathrm{d}l'. \tag{4.14}$$

Similarly, the magnetic field can be written as:

$$\begin{aligned} H_i^{tan} = &-\frac{j\omega}{k_1^2} \left(k_1^2 + \frac{\partial^2}{\partial x^2} \right) \int_l M_x(x') H_0^{(2)}(k_1|x - x'|) \, \mathrm{d}l' \\ &- \frac{j\omega}{k_2^2} \left(k_2^2 + \frac{\partial^2}{\partial x^2} \right) \int_l M_x(x') H_0^{(2)}(k_2|x - x'|) \, \mathrm{d}l' \end{aligned} \tag{4.15}$$

where the superscript *tan* indicates the tangential component of the incident magnetic field in the x-direction. Equation (4.15) represents an integro-differential equation in which the differential and integral operators on the right hand side may be interchanged, thereby obtaining:

$$\begin{aligned} H_i^{tan} = &-\frac{j\omega}{k_0^2} \int_l M_x(x') \left(k_1^2 + \frac{\partial^2}{\partial x^2} \right) H_0^{(2)}(k_1|x - x'|) \, \mathrm{d}l' \\ &- \frac{j\omega}{k_2^2} \int_l M_x(x') \left(k_2^2 + \frac{\partial^2}{\partial x^2} \right) H_0^{(2)}(k_2|x - x'|) \, \mathrm{d}l'. \end{aligned} \tag{4.16}$$

Operators with the order as in (4.16) represent *Pocklington's* integro-differential equation [98]. The second order derivative can be removed by expressing in terms of other Hankel functions through the recurrence relations [82, p. 361].

$$\frac{\mathrm{d}H_0^{(2)}(x)}{\mathrm{d}x} = -H_1^{(2)}(x) + \frac{1}{x}H_0^{(2)}(x) \tag{4.17a}$$

$$H_1^{(2)}(x) = \frac{x}{2}\left[H_0^{(2)}(x) + H_2^{(2)}(x)\right] \tag{4.17b}$$

Furthermore, a Hankel function with an argument $k_i r = k_i|x - x'|$, where $i = 1, 2$ can be differentiated by the chain-rule:

$$\begin{aligned}\frac{\partial H_0^{(2)}(k_i r)}{\partial x} &= \frac{\mathrm{d}H_0^{(2)}(k_i r)}{\mathrm{d}(k_i r)}\frac{\partial (k_i r)}{\partial x} \\ &= \frac{\mathrm{d}H_0^{(2)}(k_i r)}{\mathrm{d}(k_i r)} \times \frac{k_i(x - x')}{r}\end{aligned} \tag{4.18}$$

By differentiating (4.18) again, we obtain:

$$\frac{\partial^2 H_0^{(2)}(k_i r)}{\partial x^2} = \frac{k_i}{r}\left[H_2^{(2)}(k_i r)\frac{k_i(x - x')^2}{r} - H_1^{(2)}(k_i r)\right]. \tag{4.19}$$

The differential operator in (4.16) can now removed by applying the recurrence relations (4.17) and the expression is rewritten as:

$$\begin{aligned}\left(k_i^2 + \frac{\partial^2}{\partial x^2}\right) H_0^{(2)}(k_i r) &= \frac{k_i^2}{2}H_0^{(2)}(k_i r) + k_i^2\left[\frac{(x - x')^2}{r^2} - \frac{1}{2}\right] H_2^{(2)}(k_i r) \\ &= \frac{k_i^2}{2}H_0^{(2)}(k_i r) + k_i^2\left(\cos^2\chi - \frac{1}{2}\right) H_2^{(2)}(k_i r) \\ &= \frac{k_i^2}{2}H_0^{(2)}(k_i r) + k_i^2\cos(2\chi) H_2^{(2)}(k_i r)\end{aligned} \tag{4.20}$$

where $\cos\chi = (x - x')/r$. The magnetic field in (4.16) can be re-expressed as:

$$H_i^{tan} = \frac{-j\omega}{2} \int_l M_x(x') \left[H_0^{(2)}(k_1 r) + \cos(2\chi) H_2^{(2)}(k_1 r) + H_0^{(2)}(k_2 r) + \cos(2\chi) H_2^{(2)}(k_2 r) \right] \mathrm{d}l' \tag{4.21}$$

4.2.2.2 *TE$_z$ Polarization*

The field configuration for TE_z polarization is shown in Fig. 4.2b. Following a similar procedure and using duality principle, the field expressions in terms of currents are:

$$E_0 \sin\phi_i e^{jk_1 x \cos\phi_i} = \frac{-j\omega}{2} \int_l J_x(x') \left[H_0^{(2)}(k_1 r) + \cos(2\psi) H_2^{(2)}(k_1 r) + H_0^{(2)}(k_2 r) + \cos(2\psi) H_2^{(2)}(k_2 r) \right] \mathrm{d}l'. \tag{4.22}$$

$$\frac{E_0}{\eta_1} e^{jk_1 x \cos\phi_i} = \frac{\omega}{4} \int_l \varepsilon_1 M_z(x') H_0^{(2)}(k_1 |x - x'|) + \varepsilon_2 M_z(x') H_0^{(2)}(k_2 |x - x'|) \, \mathrm{d}l' \tag{4.23}$$

where η_1 is the characteristic impedance of external region.

4.3 Numerical Results

A method of moments (MoM) solution to compute the currents is implemented using pulse basis functions with point matching method [99] that con-

verts the integral equations into a linear system of equations,

$$\begin{bmatrix} Z_{mn} & 0 \\ 0 & Y_{mn} \end{bmatrix} \begin{bmatrix} J_n \\ M_n \end{bmatrix} = \begin{bmatrix} E_m^i \\ H_m^i \end{bmatrix} \tag{4.24}$$

where $Z_m n$ and $Y_m n$ are the impedance and admittance terms respectively. For a TM_z polarization, they are given by:

$$Z_{mn} = \frac{\omega\mu}{4} \int_{x_n}^{x_{n+1}} \left[H_0^{(2)}(k_1|x_m - x_n|) + H_0^{(2)}(k_2|x_m - x_n|) \right] \mathrm{d}x' \tag{4.25a}$$

$$\begin{aligned} Y_{mn} = \frac{\omega}{8} \int_{x_n}^{x_{n+1}} \Big[& \varepsilon_1 H_0^{(2)}(k_1|x_m - x_n|) + \varepsilon_2 H_0^{(2)}(k_2|x_m - x_n|) + \\ & \varepsilon_1 H_2^{(2)}(k_2|x_m - x_n|) + \varepsilon_2 H_2^{(2)}(k_2|x_m - x_n|) \Big] \mathrm{d}x'. \end{aligned} \tag{4.25b}$$

The integrals in (4.25) become singular for $(m = n)$ for which small argument approximations of the Hankel function are used to estimate the integral,

$$Z_{nn} = -\frac{\omega\mu\Delta}{8} \left\{ 2 - \frac{j}{\pi} \left[6 + 2\ln\left(\frac{\mathrm{e}^{\gamma} k_1 \Delta}{4e}\right) + \frac{16}{(k_1\Delta)^2} + 2\ln\left(\frac{\mathrm{e}^{\gamma} k_2 \Delta}{4e}\right) + \frac{16}{(k_1\Delta)^2} \right] \right\} \tag{4.26a}$$

$$Y_{nn} = -\frac{\omega\mu\Delta}{4} \left\{ 2 - \frac{2j}{\pi} \ln\left(\frac{\mathrm{e}^{\gamma} k_1 \Delta}{4e}\right) - \frac{2j}{\pi} \ln\left(\frac{\mathrm{e}^{\gamma} k_2 \Delta}{4e}\right) \right\} \tag{4.26b}$$

4.3.1 Current Distribution

Figure. 4.3a shows the absolute value of the tangential surface electric current on a TM_z polarized plate of length 3λ at edge-on ($\phi_i = \pi$). Gallium Arsenide (GaAs) and Strontium Titanate ($SrTiO_3$) sheets are considered at terahertz fre-

quencies where material data has been taken from measurements in [100] and [101] respectively, and the results are compared with a PEC plate of same length [102].

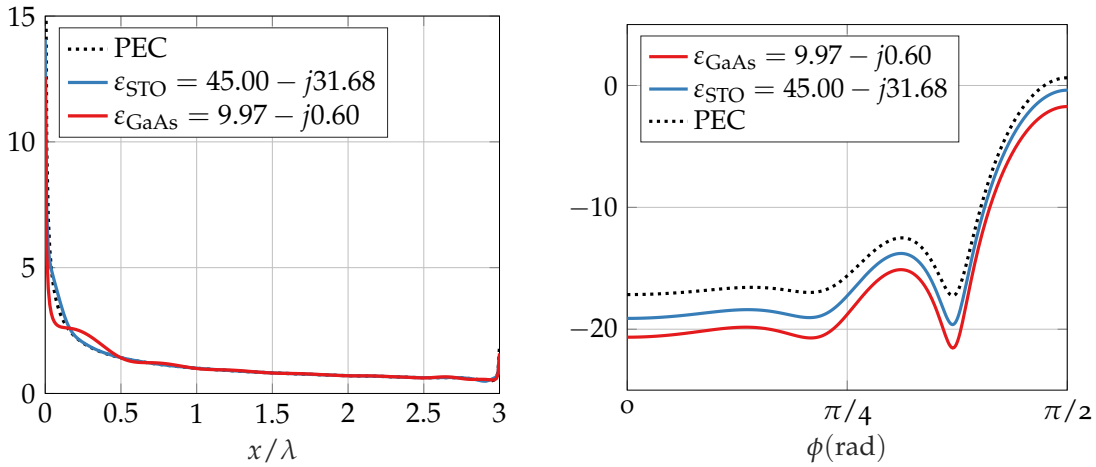


Figure 4.3: (a) Current distributions on a 3λ plate at edge-on incidence, (b) backscattered fields from different sheets of length 1.25λ

4.3.2 Far-field

The scattered electric field in the far-zone can be expressed by normalizing the large argument approximation of Hankel functions as:

$$\lim_{k_1|\boldsymbol{\rho}-\boldsymbol{\rho}'|\to\infty} E_z(\boldsymbol{\rho}) \simeq \int_0^L J_z(x')\mathrm{e}^{jk_1x'\cos(\phi_i)}\,\mathrm{d}x' \tag{4.27}$$

where ϕ_i is the angle of incidence. The results are shown in Fig. 4.3b.

4.3.2.1 *Comparison with Other Techniques*

Here we compare results our scheme with other techniques that have been used in the context of thin sheets. First, the far-field patterns of TM_z polarized dielectric sheet of permittivity $\varepsilon = 4$ are compared with results in [103] computed using a volume integral equation (VIE) technique collapsed to a surface. The length of the dielectric rod is taken as 2.5λ and the radar cross-sections (RCS) are shown in Fig. 4.4a. For the same dielectric structure, the RCS for TE_z is plotted in

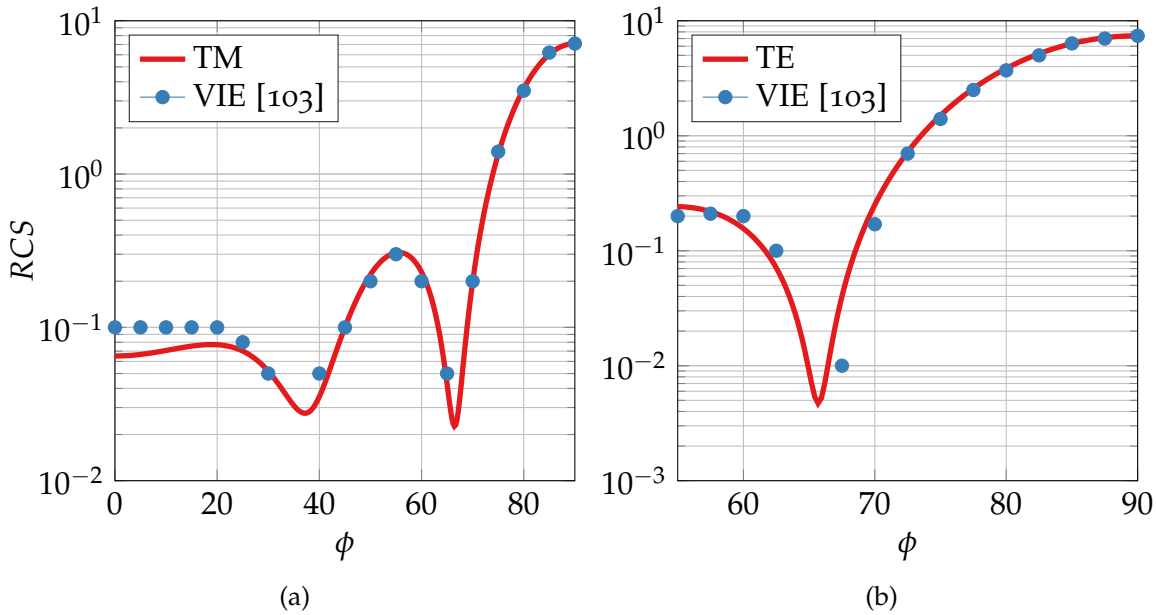


Figure 4.4: Computed RCS of a 2.5λ dielectric rod of permittivity $\varepsilon = 4$ (a) TM_z, (b) TE_z

Fig. 4.4b. The disparity in the results near edge-on incident angles is attributed to the fact that we consider an infinitesimally thin sheet whereas in the reference values, a finite dielectric rod thickness of 0.05λ was assumed.

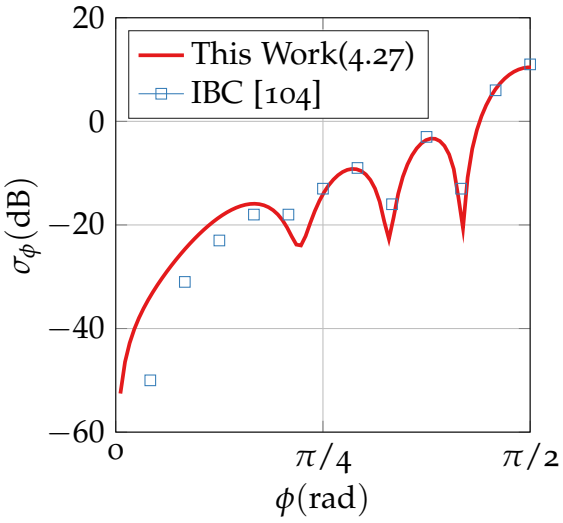


Figure 4.5: Comparsion of RCS of a 2λ dielectric sheet having $\varepsilon = 4$ with a resistive sheet model

Next we consider a TE_z polarized dielectric rod of length 2λ and compare our results with the ones presented in [104] that were computed using a resistive sheet model which is derived from the impedance boundary conditions (IBC) through the expression:

$$E^i = Z\, J_s(x') + \frac{\omega\mu}{4} \int_l J_s(x') H_0^{(2)}(k_2|x - x'|)\, dx'. \tag{4.28}$$

The results are shown in Fig. 4.5. Again, the diverging results away from normal incidence are due to the difference in sheet thickness where it was taken as 0.1λ in the reference and zero in our computations. We present a new class of surface integral equations for infinitesimally thin dielectric sheets based on surface equivalence theorem. The electromagnetic response to external radiation is investigated. Results are shown for the current induced and far-zone response,

and compared with other techniques that are used to simulate thin sheets.

5. SURFACE WAVES EXTRACTION OF A CONDUCTIVE SHEET EMBEDDED IN A MULTILAYER ENVIRONMENT

In this section, we compute the dispersion diagrams of an infinitesimally thin conductive that lies embedded in a stack of semiconductors. The existence of surface wave propagation is investigated by searching for poles of an equivalent transmission line Green function determined by the multilayer structure. A numerical root-finding routine is presented that permits complex-valued solutions and avoids singularities in the complex plane.

5.1 Theory

5.1.1 Equivalent TL Network of Semiconductor Heterostructures

Modern field-effect transistors are engineered by epitaxially growing multiple layers of semiconducting materials with slightly different dielectric properties and band-gap energies. Of particular interest to the microwave and radiofrequency (RF) community are the group III-V materials in which extra-ordinary electromagnetic interfacial phenomena is observed in the terahertz frequency domain, mainly due to formation of a two-dimensional electron gas (2DEG) at the interface. To develop an understanding of the unusual electromagnetic properties of what are commonly known as semiconductor heterostructures, we construct an equivalent transmission model of the layered structure using the theory discussed in the previous section.

An illustration of a multilayer setup with a heterostructure near the top of the transistor substrate, through which modern devices like the high electron mobility transistor operate, is shown in Fig. 5.1a. Although a number of other material layers lie underneath the heterostructure that are essential to the fabri-

cation process, primarily to strengthen the epitaxial structure, only the top layers are conducive to an electric field. Moreover, since the widths of the layers in the heterostructure are much smaller than the rest of the layers in the substrate stack, we assume that the heterostructure is unbounded from the bottom.

The 2DEG which is only a few atoms wide in thickness, is a highly conductive region with an abundance of free electrons, which can be expressed in terms of a surface conductivity [100],

$$\sigma_s = \frac{N_s e^2 \tau}{m^*} \frac{1}{1 + j\omega\tau}, \tag{5.1}$$

where N_s is the free electron concentration at the interface, e is the electron charge, m^* is the effective electron mass in the heterostructure, τ is the scattering time of electrons, and ω is the angular frequency. For a wave propagating along the interface, the electron scattering time varies greatly with temperature, and it determines the attenuation factor. At room temperature, the thermal vibrations of electrons result in a very small value of τ which is of the order of femtoseconds (1×10^{-15} s), leading to a high-loss scenario. On the other hand, when the heterostructure is cooled to temperatures close to the boiling point of helium (4.2 K), thermal vibrations are effectively removed and the losses are consequently minimized.

In a transmission line analogue, the equivalent representation of a conductive sheet such as a 2DEG is a shunt admittance as shown in Fig. 5.1b. When a voltage bias is applied across the source and drain terminals of the transistor, a current starts to flow laterally, inside the 2DEG. The presence of a gate terminal above or below the interface allows the electron concentration to be modified by the gate voltage. Therefore, instead of an independent current source in the equivalent

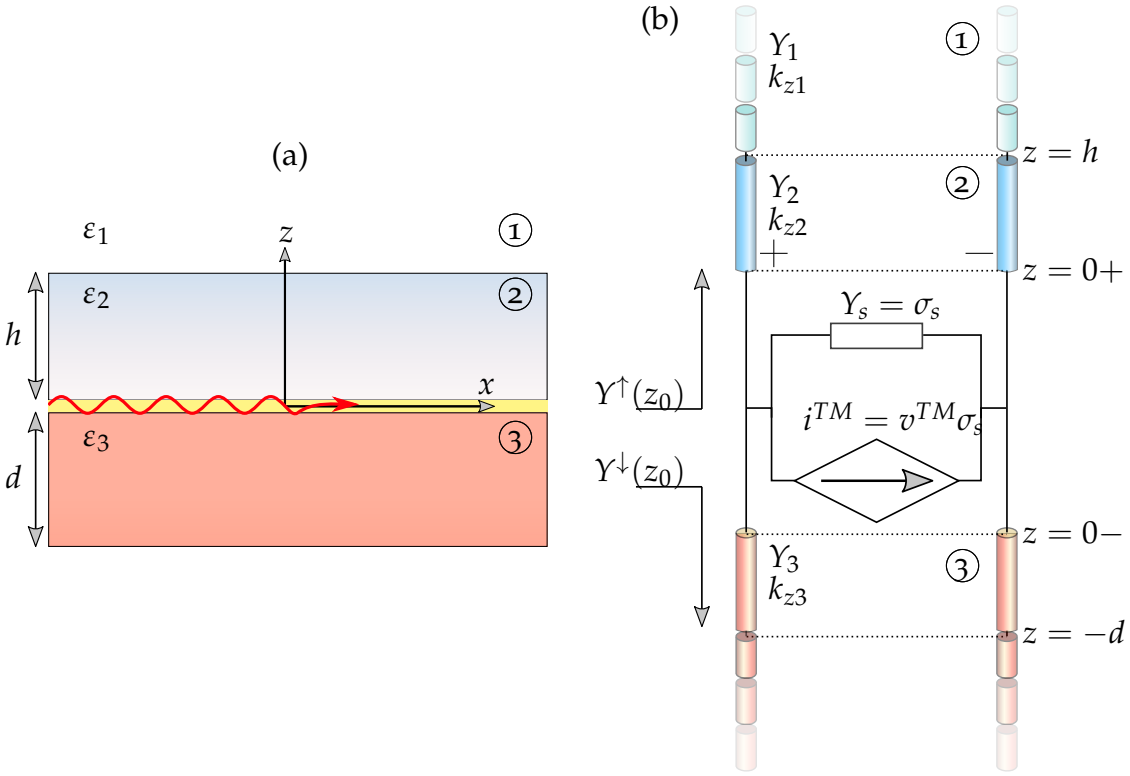


Figure 5.1: (a) Multilayer structure typically found in a high electron mobility transistor, (b) equivalent transmission line network

transmission line network, a voltage-controlled current source is placed as indicated in Fig. 5.1. We determine the dispersion relation in a 2DEG region which lies embedded in a semiconductor stack. As stated earlier, the active state of the transistor in which a current flows through the 2DEG corresponds to a TM-mode transmission line network. It must also be mentioned that a similar TL analogue can be obtained if the physical structure is excited by a TM-polarized plane wave. The dispersion relation can be written using the transverse resonance method that requires the total admittance as seen from the 2DEG (at z = 0) to be zero [78, 105],

$$Y^{\uparrow}(z_0) + Y^{\downarrow}(z_0) + Y_{\sigma} = 0. \tag{5.2}$$

Here, $Y^{\uparrow}(z_0)$ and $Y^{\downarrow}(z_0)$ are the up- and down-looking TL admittances from the 2DEG located at $z = 0$ and $Y_\sigma = \sigma_s$ from (5.1).

A nontrivial solution of (5.2) in terms of the propagation constant k_ρ yields the surface wave modes of the structure in the complex k-plane.

5.2 Surface Waves Extraction Technique

We seek to solve (5.3) in the complex z-plane to obtain the zeros, z_k in a given search region $\mathbb{C}$. For a multilayer problem such as the one shown in Fig. 5.1a, f represents the dispersion relation (5.2) which, in fact is the denominator of the transmission line Green function (TLGF). From hereon, k_ρ is replaced by z for notational simplicity. Depending on the location of the zeros of the function f in the complex plane, z_k's can be termed as the surface wave modes of the multilayer or synonymously, poles of the TLGF.

$$f(k_\rho) = 0 \tag{5.3}$$

Although the problem at first, may appear to be a relatively simple root-search in which readily available and popular iterative algorithms like the Newton or Halley's method can be applied, complication arises due to the presence of complex-valued roots. Furthermore, the convergence of the aforementioned methods is highly subject to an initial guess that must lie close to the actual root. As an example, basins of attraction, which illustrate the convergence rate of an initial guess, are shown in Fig. 5.2 for Newton's and Halley's methods which are two of the most popular iterative root-finding methods. In Fig. 5.2, the actual roots of a polynomial, $z^{10} - 1$ are depicted by circle marks. The regions are color-coded for each root where a large area of a particular color indicates that if the initial guess is located in that region, convergence to the respective root will be

achieved quickly. On the other hand, if the initial guess lies in the fractal region, the convergence is very slow. Another observation through which Newton's and Halley's methods can be compared is that the basins of attraction of Newton's method, shown in Fig. 5.2a have a denser fractal region due to its quadratic complexity of convergence. On the contrary, the Halley's method, which is cubic, the fractals are comparatively sparse as shown in Fig. 5.2b.

For a dispersion relation (5.2), which is defined by a transcendental equation, presence of singularities, particularly in the vicinity of an actual solution may leave an iterative method completely divergent or yet churn out a bogus and unreliable answer.

Argument principle method (APM) is a robust, complex root-finding algorithm in which convergence is guaranteed within a specified region without supplying any initial guess [106, 107, 108, 109, 110, 111, 112]. It requires the function to be analytic in the specified search region. A pictorial representation of this technique is shown in Fig. 5.3 where the roots inside a specified region Γ in the complex plane are computed by approximating the given function with a polynomial. In case, the number of zeros inside Γ exceed a pre-allocated value N_r, the search region is subdivided into smaller regions. Moreover, the function f must be non-zero at the boundary of Γ. The following sections briefly describe each step of the method.

5.2.1 Counting the Zeros

To develop an efficient method of locating the zeros, we assume that the function f is analytic, which implies that it is differentiable and free from any singularities. It is also required that f is non-zero at the boundary, Γ of the region $\mathbb{C}$. The search region is set as rectangular, mainly because of ease in

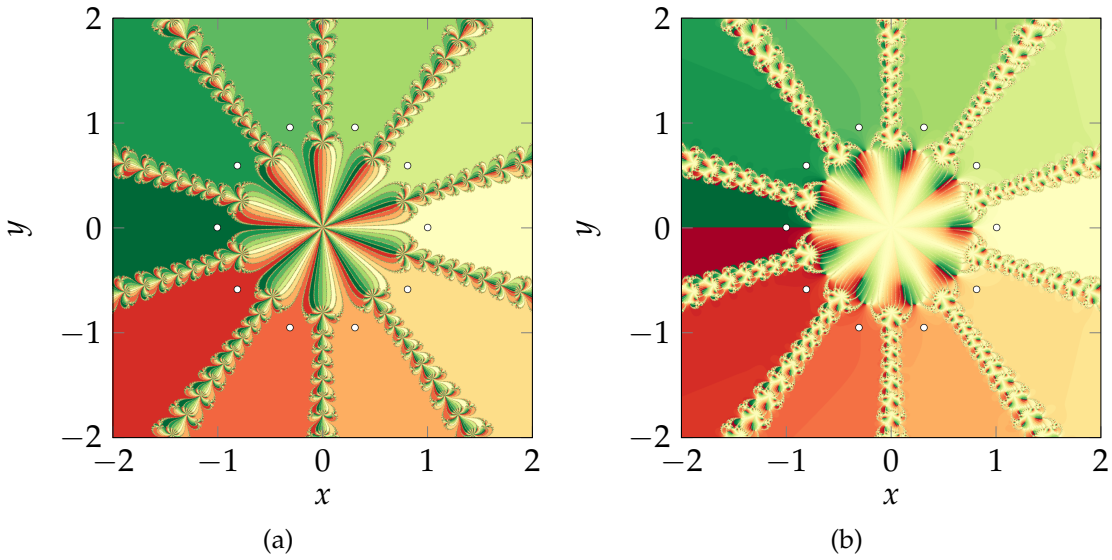


Figure 5.2: Complex plane basin of attraction while solving $z^{10} - 1 = 0$ using: (a) Newton's method (c) Halley's method

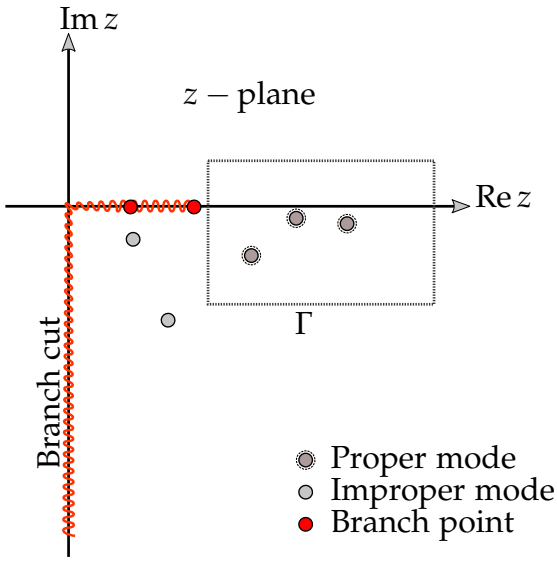


Figure 5.3: Illustration of the root-finding routine in the presence of branch point singularities. Only proper modes are sought within a rectangular contour Γ

programming it. The number of zeros, N inside a region with a boundary Γ can be found using the Cauchy's integral theorem [113, 106, pg. 71],

$$N = \frac{1}{2\pi j} \oint_{\Gamma} \frac{f'(z)}{f(z)} \, \mathrm{d}z \tag{5.4}$$

where the integrand is a logarithmic derivative of the function. The analytical derivative of functions such as found in dispersion relations like (5.2) are seldom readily-available and it is generally cumbersome to compute. In this section, we use approximate the derivative by a finite difference formula,

$$f'(z) = \frac{f(z+h) - f(z-h)}{2h}, \tag{5.5}$$

where $h \sim \sqrt{\epsilon_m}$ and $\epsilon_m = 2.220 \times 10^{-16}$ is the double-precision machine accuracy [114, pg. 230].

In addition to (5.4), a derivative-free form of the argument principle [107, 111], which may be preferred in certain cases is,

$$N = \frac{1}{2\pi} \oint_{\Gamma} \mathrm{d}\{\arg f(k_\rho)\}, \tag{5.6}$$

The value of both the integrals, i.e., (5.4) and (5.6) is an integer which means that whenever a zero is encountered that appears as a pole in the integrand, the resulting integral is incremented by a factor of 2π by Cauchy's residue theorem. The contour integration in (5.4) and (5.6) is computed in a counter-clockwise manner, using an adaptive Gauss-Konrod quadrature (MATLAB's *quadgk* routine) [115]. The contour can be conveniently set up using the 'waypoints' parameter of the routine.

5.2.2 Locating the Zeros

In this section, we use the derivative-free approach (5.6) to compute the number of zeros and further on. After predicting the number of zeros, the function f is then approximated by an associated *formal orthogonal polynomial* (FOP), $\mathcal{P}$ of degree N, equal to the number of zeros. A necessary condition during approximation is that $\mathcal{P}$ must have the same roots, z_k where $k = 1, 2, ..., N$ as f. Furthermore, as stated earlier N must be less than N_r to avoid a too-high order polynomial for which the root-finding procedure would be very slow.

The polynomial $\mathcal{P}$ can be represented in a Lagrange form:

$$\begin{aligned} \mathcal{P}_N(z) &= \prod_{k=1}^{N} (z - z_k) \\ &= z^N + \sigma_{N-1}\, z^{N-1} + \cdots + \sigma_1\, z + \sigma_0. \end{aligned} \tag{5.7}$$

To find the coefficients, σ_k we first define a symmetric bilinear operator similar to a dot product [116, 111],

$$\langle \phi, \psi \rangle = \frac{1}{2\pi j} \oint_{\Gamma} \phi(z)\, \psi(z) \frac{1}{f(z)}\, \mathrm{d}z \tag{5.8}$$

where ϕ and ψ are polynomials of z. Next, we consider a set of integrals that are termed as the *Newton moments*,

$$s_p = \langle 1, z^p \rangle \frac{1}{2\pi j} \oint_{\Gamma} z^p \frac{1}{f(z)}\, \mathrm{d}z. \tag{5.9}$$

which are then used as elements to construct Hankel matrices of size N,

$$\mathbf{H} = \begin{bmatrix} s_1 & s_2 & \cdots & s_k \\ s_2 & s_3 & \cdots & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_k & s_{k+1} & \cdots & s_{2k} \end{bmatrix} \tag{5.10}$$

and

$$\mathbf{H}^{<} = \begin{bmatrix} s_0 & s_1 & \cdots & s_{k-1} \\ s_1 & s_2 & \cdots & s_k \\ \vdots & \vdots & \ddots & \vdots \\ s_{k-1} & s_k & \cdots & s_{2k-2} \end{bmatrix}, \tag{5.11}$$

in which the secondary diagonal terms are the same.

The polynomial $\mathcal{P}$ in (5.7) is orthogonalized by enforcing the condition,

$$\left\langle z^k, \mathcal{P}_N(z) \right\rangle, \quad \text{where } k = 0, 1, \ldots, N-1. \tag{5.12}$$

Because the operator $\langle \cdot, \cdot \rangle$ defined in (5.8) is not a dot product in true sense, that is why the term *formal* is included in the definition of the polynomial [116].

The polynomial in (5.7) can also be equivalently represented by a *companion* matrix, A as,

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 \\ -\sigma_0 & -\sigma_1 & -\sigma_2 & \cdots & \sigma_{N-1} \end{bmatrix}, \tag{5.13}$$

and $\mathcal{P}_n(z)$ can be obtained by setting its characteristic equation [117],

$$\det(A - \sigma_k I) = P(z), \tag{5.14}$$

where the eigenvalues are given by σ_k for the eigenvector z. The eigenvalue value problem of (5.14) can not be solved since the coefficients σ_k's, where $k = 0, 1, 2, \ldots, N-1$, are not known yet. However, we can express σ_k's in terms of the *Newton moments*, s_k's through the *Newton identities*:

$$\begin{aligned} & s_1 + \sigma_1 = 0 \\ & s_2 + s_1\sigma_1 + 2\sigma_2 = 0 \\ & \vdots \\ & s_N + s_{N-1}\sigma_1 + ... + s_1\sigma_{N-1} + s_0\sigma_N = 0. \end{aligned} \tag{5.15}$$

where we note that $s_0 = N$. We therefore, construct an equivalent eigenvalue problem using (5.10), (5.11) and (5.15) [111],

$$\left(\mathbf{H} - \lambda_k \mathbf{H}^{<}\right) \mathbf{z} = 0 \tag{5.16}$$

where the column vector $\mathbf{z}$ is an eigenvector and λ_k's are the eigenvalues of the matrix pencil $(\mathbf{H} - \mathbf{H}^{<})$, which correspond to the roots of $\mathcal{P}_n(z)$. In MATLAB, the routine *eig* is used to solve (5.16).

5.2.3 Refining the Roots

The accuracy of the roots obtained from the eigenvalues, λ_k of (5.16) is not always high. However, λ_k's is an excellent initial guess for any iterative root-search routine from the class of Householder's methods. We choose the *Newton*

method [114] having quadratic convergence and the iteration formula:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{5.17}$$

with $f'(x)$ is the first order derivative approximated by a finite difference formula (5.5). In general, the roots z_k's are complex, therefore, the iteration (5.17) needs to be performed on both the real and imaginary parts simultaneously.

5.2.4 Avoiding Singularities

As mentioned earlier, the argument principle method requires the function to be analytic without any singularities. In this section, we explicitly deal with the dispersion relation (5.2) which is the denominator expression of the transmission line Green function. Therefore, branch point singularities only occur due to the presence of multivalued square roots that appear in the definition of the propagation constant, $k_{z,i} = (k_i^2 - k_x^2)^{1/2}$ in which $i = 1, \ldots, N$ and, N being the number of layers, k_i is the wavenumber of the i-th layer and k_x is the propagation constant in the lateral direction. The branch points exist only when the top and bottom layers are unbounded. The dispersion relation shows an even dependence on k_{zi} due to the layers that are bounded, hence they do not contribute any branch points [78, Section 5.3a]. In the worst case where the top and bottom layers are open and different from each other, two branch points exist in the complex plane that must be either removed or avoided in order to make the function analytic in the search region. Additionally, branch cuts emerge out of the branch points and there must not be any intersection with the contour of the search region.

To remove the branch point singularities, the complex z-plane is typically mapped to another plane ζ, by a transformation that unfolds the Riemann surface.

For example, a trigonometric transformation such as [118, 119],

$$\begin{aligned} z &\mapsto \zeta \\ \zeta &= \cos(z), \end{aligned} \tag{5.18}$$

removes the branch point due to the multivalued propagation constant, k_{zi}, where $i = 1, N$ and makes the dispersion relation (5.2) analytic in the ζ-plane.

In this section, our goal is to find the surface waves from the dispersion relation. We only seek the *proper modes*, which are the roots of (5.2) that lie to the right of the branch point corresponding to $\max(k_{z1}, k_{zN})$. In order to satisfy the Sommerfeld radiation condition which requires all fields to decay to zero at infinity, the proper sheet of the Riemann surface of the square-root function needs to be selected. This is accomplished by enforcing that the imaginary part of k_{zi} that contributes a branch point, is negative everywhere.

$$\mathrm{Im}(k_{zi}) < 0 \tag{5.19}$$

5.3 Results

5.3.1 MIM Waveguide

We first compute the modes of a metal-insulator-metal (MIM) waveguide described in [120]. An air-filled region of thickness $d = \lambda/4$, where $\lambda = 1550\,\mathrm{nm}$ is sandwiched between two metal layers having permittivity ε_m. We first consider a lossless case in which ε_m has a value of -143.4967, because of which the complex plane is essentially free from any branch-point singularities. For a TM case the

dispersion relation is written as:

$$D = 1 - \left(\frac{Z_m - Z_2}{Z_m + Z_2} \right)^2 \times \mathrm{e}^{(-2jk_{z2}d)}, \tag{5.20}$$

where $Z_m = k_{zm}/(\omega\varepsilon_m\varepsilon_0)$ and $Z_2 = k_{z2}/\omega\varepsilon_0$ are the modal impedances of the metal and air-filled layers respectively. The TM modes of (5.20) normalized to the free-space propagation constant are shown in Fig. 5.4a. The values of the roots obtained are then compared with the solutions discussed in [120], in which only the even modes were computed. However, as listed in Table. 5.1, we compute first nine even and odd modes.

Mode	This method (5.20)	Kocabaş et. al. [120]
TM_0	$12.02 - j0.004\,432$	12.03
TM_1	$11.85 - j0.008\,866$	
TM_2	$11.34 - j0.008\,868$	11.34
TM_3	$10.42 - j0.008\,867$	
TM_4	$8.977 - j0.008\,867$	8.981
TM_5	$6.681 - j0.008\,867$	
TM_6	$0.8324 - j0.008\,867$	0.7137
TM_7	$0. - j7.149$	
TM_8	$0. - j10.53$	$0.008\,873 - j10.56$

Table 5.1: TM modes of a lossless MIM waveguide

Next, we consider the lossy case where the permittivity of metal layers is $\varepsilon_m = -143.5 + j9.517$ and the compared results are shown in Fig. 5.4a and listed in Table. 5.2.

Mode	This method (5.20)	Kocabaş et. al. [120]
TM_0	$12.03 - j0.4001$	$12.03 - j0.3954$
TM_1	$11.86 - j0.4100$	
TM_2	$11.35 - j0.4283$	$11.35 - j0.4189$
TM_3	$10.43 - j0.4651$	
TM_4	$8.994 - j0.5380$	$8.996 - j0.5289$
TM_5	$6.720 - j0.7170$	
TM_6	$2.267 - j2.108$	$2.248 - j2.125$
TM_7	$0.6612 - j7.188$	
TM_8	$0.4504 - j10.54$	$0.4591 - j10.57$

Table 5.2: TM modes of a lossy MIM waveguide

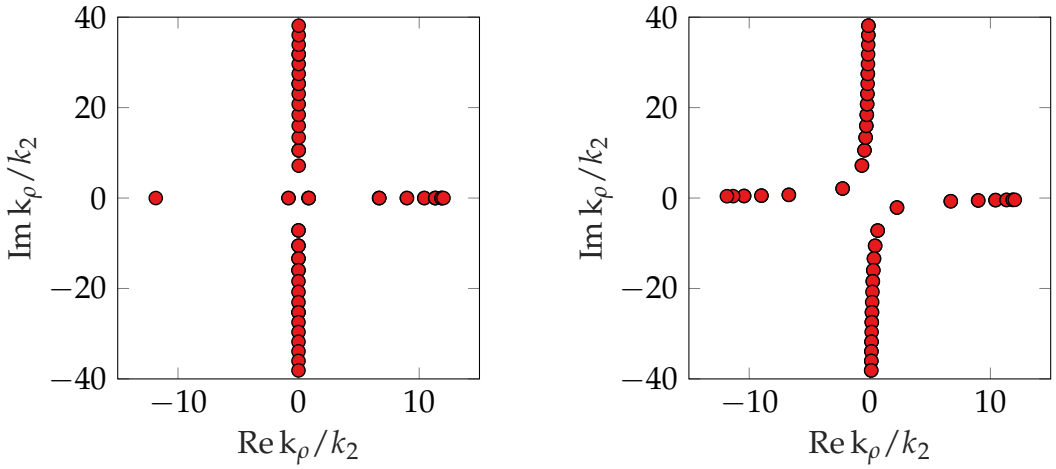


Figure 5.4: TM modes of a MIM waveguide: (a) lossless case, ($\varepsilon_m = -143.5$) (b) lossy case, ($\varepsilon_m = -143.5 + j9.517$)

5.3.2 Embedded Conductive Sheets

Here we plot the dispersion curves of a 2DEG which is considered as a thin sheet embedded in a semiconductor heterostructure. We look at three types of structures,

1. Gated 2DEG
2. Ungated 2DEG
3. Backgated 2DEG,

that are classified, based on the presence of a perfectly conducting covering layer either at the top or bottom. The three types are shown in Fig. 5.5.

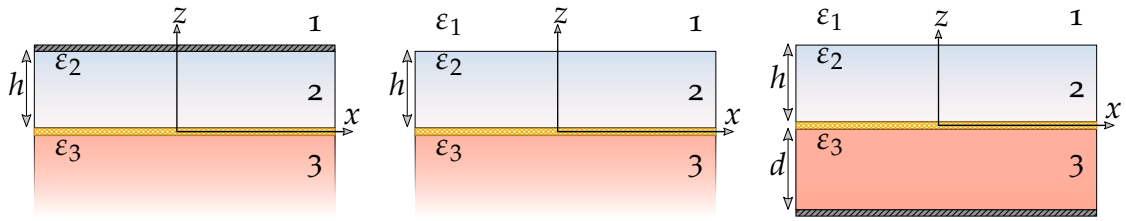


Figure 5.5: 2DEG embedded in a semiconductor heterostructure at $z = 0$. (a) Gated, (b) backgated, and (c) ungated.

We consider GaN/AlGaN (gallium nitride / aluminum gallium nitride) heterostructure with Al mole-fraction of $x = 0.2$. With respect to Fig. 5.5, Layer 1 is composed of GaN which is the substrate, considered open in the gated and ungated cases, while it has a thickness of $d = 50\,\text{nm}$ in the backgated case. AlGaN forms the layer marked 2, and in all the three cases, a thickness, $h = 20\,\text{nm}$ is assumed. The permittivity of both the layers is $\varepsilon_r = 9.5$ owing to the small

aluminum mole fraction [92]. The electron concentration of the 2DEG region is $N_s = 5 \times 10^{13}\,\mathrm{cm}^{-2}$ and scattering time τ of 114 ps corresponding to a temperature of 3 K is assumed. The surface conductivity is then computed using (5.1). When solving for the dispersion relation (5.2), we use the respective expressions listed in Table. 5.3 for each case. The dispersion curves for each structure type

Gated	Backgated	Ungated
$Y^{\uparrow} = -Y_2 \coth(k_{z2}h)$	$Y^{\uparrow} = Y_2 \frac{1-\Gamma^{\uparrow}}{1+\Gamma^{\uparrow}}$	$Y^{\uparrow} = Y_2 \frac{1-\Gamma^{\uparrow}}{1+\Gamma^{\uparrow}}$
	$\Gamma^{\uparrow} = \frac{Y_2-Y_1}{Y_2+Y_1} \mathrm{e}^{-2jk_{z2}h}$	$\Gamma^{\uparrow} = \frac{Y_2-Y_1}{Y_2+Y_1} \mathrm{e}^{-2jk_{z2}h}$
$Y^{\downarrow} = Y_3$	$Y^{\downarrow} = -Y_3 \coth(k_{z3}d)$	$Y^{\downarrow} = Y_3$

Table 5.3: Equivalent upward- and down-looking admittances

at varying temperatures and electron concentrations are plotted in Figs. 5.6 and 5.7. To observe the effect of temperature, we now plot the dispersion curves at room temperature ($T = 295\,\mathrm{K}$) which corresponds to a scattering time, τ of 114 fs. The conducting layer in the gated and back-gated cases depicts the gate terminal which is voltage biased in a normal transistor operation. By varying the voltage, we can subsequently modify the electron density in the 2DEG channel. In Figs. 5.8 and 5.9, the dispersion curves of gated and backgated structures are plotted for $N_s = 1 \times 10^{12}\,\mathrm{cm}^{-2}$ at 3 and 295 K respectively. In the results shown, the lowest order TM_0 is plotted in all cases. By comparing the plots in Fig. 5.6 with the corresponding plots in Fig. 5.6, it is apparent that by increasing the temperature from liquid helium state tot room temperature, loss expressed in terms of the

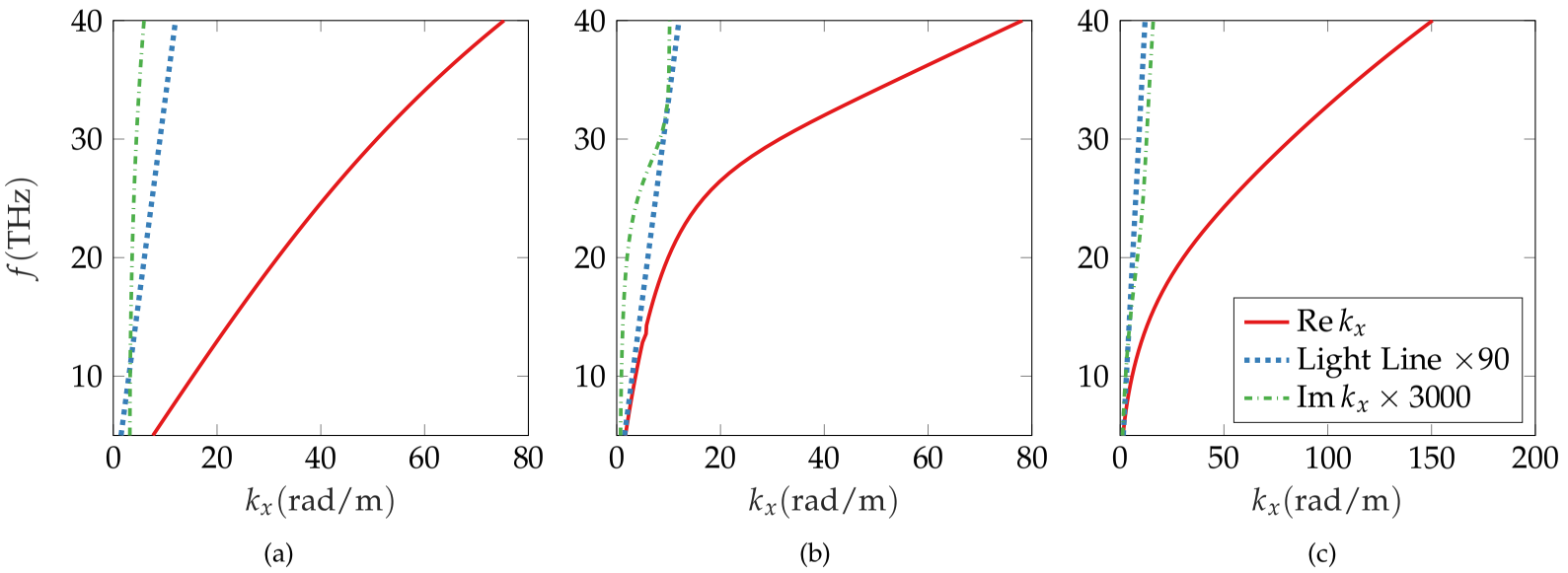


Figure 5.6: Dispersion curves plotted for a GaN/AlGaN heterostructure with $N_s = 5 \times 10^{13}\,\mathrm{cm}^{-2}$ at 3 K (a) Gated (b) Backgated (c) Ungated

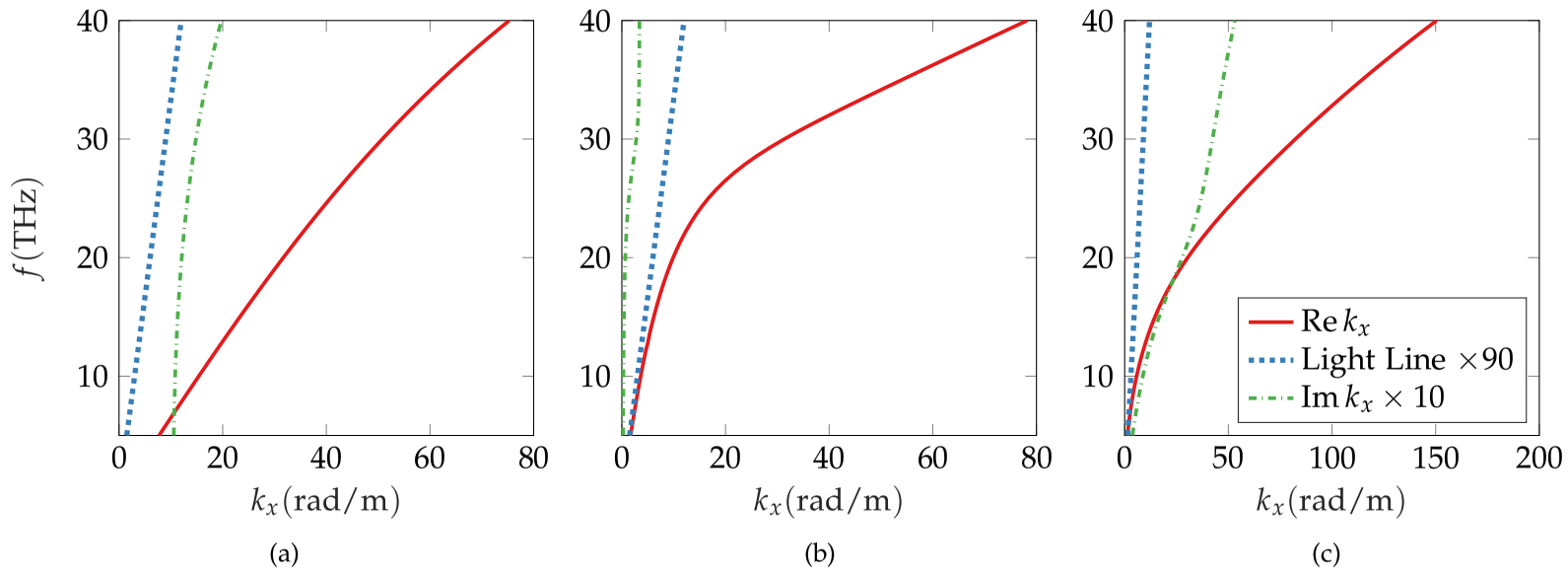
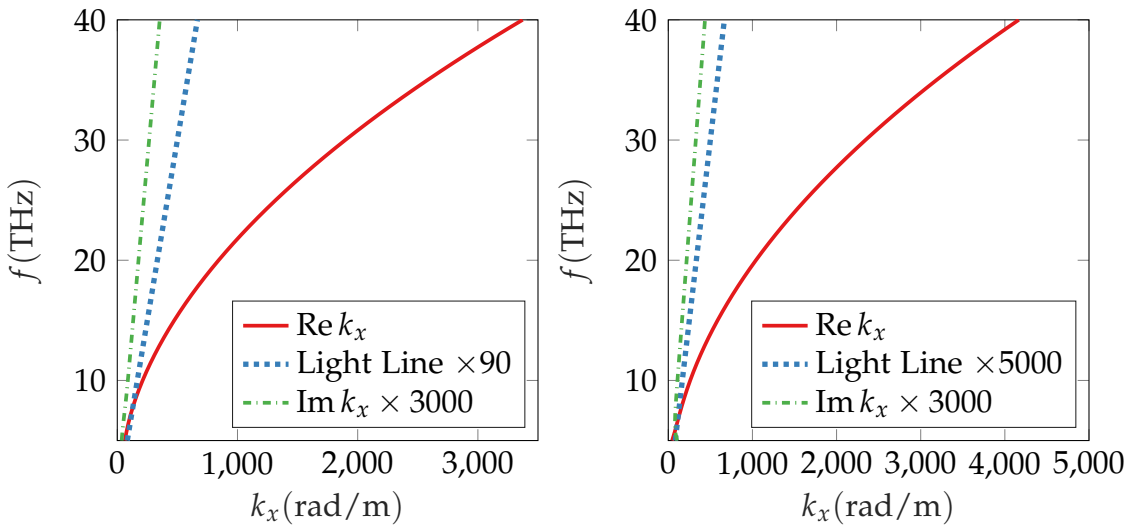
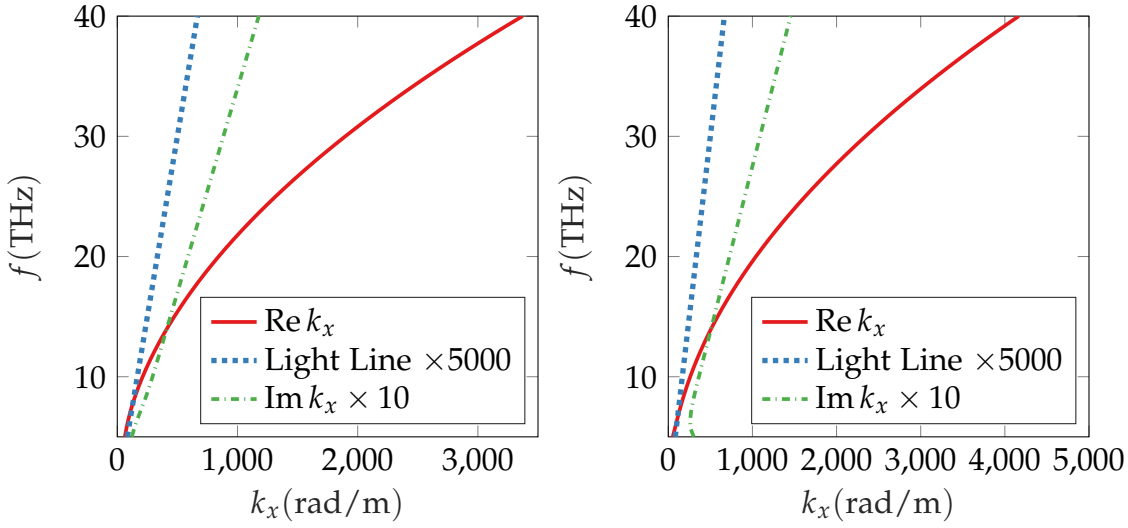


Figure 5.7: Dispersion curves plotted for a GaN/AlGaN heterostructure with $N_s = 5 \times 10^{13}\,\mathrm{cm}^{-2}$ at 295 K. (a) Gated, (b) backgated, and (c) ungated.



(a) (b)

Figure 5.8: Dispersion curves plotted for a GaN/AlGaN heterostructure with $N_s = 1 \times 10^{12}\,\mathrm{cm}^{-2}$ at 3 K. (a) Gated and (b) backgated.



(a) (b)

Figure 5.9: Dispersion curves plotted for a GaN/AlGaN heterostructure with $N_s = 1 \times 10^{12}\,\mathrm{cm}^{-2}$ at 295 K. (a) Gated, and (b) backgated.

imaginary part of the propagation constant is introduced. The imaginary part and the light line are enhanced for illustration.

The propagation constant belonging to the gated region exhibits a linear relationship [121, 122] as shown in Figs. 5.6a and 5.7a. Whereas a parabolic type dispersion curve is observed for the ungated region in Figs. 5.6c and 5.7c. For the backgated structure, the plot is initially linear for lower frequencies, followed by a parabolic region. The behavior is shown in Figs. 5.6b and 5.7b. Such a dispersion is generally attributed to surface plasmon polaritons existing at a dielectric-metal interface [123].

Figures 5.8 and 5.9 show the plots of gated and backgated dispersion curves in which the 2DEG region is tuned to a lower electron concentration of $N_s = 1 \times 10^{12}\,\mathrm{cm}^{-2}$. Reduction of N_s by more than an order of magnitude results in a marked increase of the propagation constant [124].

5.4 Conclusion

The dispersion relations of various configurations of a GaN/AlGaN heterostructure were numerically solved. The computed dispersion curves clearly show that the propagation constant in the 2DEG region, which is embedded in the multilayer, is much larger than free-space wavenumber. This translates to the ability to support subwavelength phenomenon. It was also shown the efficiency of such structures is greatly reduced at high temperatures due to increased amount of loss.

6. NANOSCOPY USING A SEMICONDUCTOR HETEROSTRUCTURE AS THE SAMPLE STAGE

A special structured illumination microscopy scheme using a two dimensional electron gas as the sample stage is proposed. Terahertz plasma waves generated by a current-driven instability illuminate the sample. Concurrently, a plane wave is used to shift the plasmonic pattern needed to expand the observable range of spatial frequencies. Full coverage of the spatial frequency regime is obtained by tuning the plasma waves through gate voltage control. Hence, it is possible to reconstruct an image with resolution up to two orders of magnitude beyond the diffraction limit. Due to the linear nature of the technique, only a weak illumination signal is required, therefore minimizing the likelihood of sample damage due to thermal effects.

6.1 Introduction

In conventional wide-field fluorescent microscopy, a sample is uniformly illuminated by a beam of light, and the resulting fluorescence is observed in the far-field through the objective lens. The uniform intensity of the illumination along the sample fundamentally restricts the resolution of the system to half the wavelength of light due to the Abbe diffraction limit. With ever growing need to image miniscule objects especially in life sciences, modern microscopy techniques such as confocal and linear structured illumination microscopy (SIM) use spatially non-uniform sources of light to illuminate the sample. This non-uniformity is directly attributed to obtaining resolution that can be extended beyond the diffraction limit by a factor of 2 [125, 126]. Use of pinholes in confocal microscopy makes the technique highly inefficient as a significant part of light

is discarded that may leave weakly fluorescent objects undetectable. Structured Illumination microscopy is a wide-field technique in which a fine illumination pattern such as a sinusoidal standing wave is used to generate *Moiré fringes* in the observed image [127, 128]. The high frequency content is mathematically reconstructed from a series of images acquired by shifting the pattern. Using a non-linear version of SIM, theoretically unlimited resolution can be achieved [129]. However, high levels of illumination intensity are required, subjecting the sample to significant damage due to thermal effects.

Resolution beyond the classical diffraction by a factor greater than 2 can be realized when an object is illuminated by surface waves, that are electromagnetic waves found at a material interface where one of the media possesses a complex dielectric function with a negative real part. The wavelength and phase velocity of the surface waves are much smaller than the homogeneous waves at the same frequency in a given medium. Using a grating structure to generate the surface waves with sub-diffraction features, a resolution with magnification factor of 20 was first demonstrated at optical frequencies by Nassenstein [130]. Recently, a plasmonic structured illumination microscopy (PSIM) technique was proposed in which a sample was excited by surface plasmons existing at a metal-dielectric interface in the optical frequency range [131]. In another scheme, it was shown that a hundred-fold resolution enhancement is possible using mid-infrared graphene plasmons [132, 133].

Current-drive plasma instabilities have mainly been studied in the context of ionized gases [134]. An analogous activity leads to generation of plasma waves in solid-state devices [3] that has led to many interesting applications in the far-infrared frequency region [135, 136]. A two-dimensional electron gas (2DEG) is formed at the interface of epitaxially grown semiconductors that acts as a tran-

sistor channel. In addition to the unusually high electron mobility, free-electron densities comparable to metals are observed without any intentional doping in the channel. Plasma waves originating in the field-effect transistor electron channel, which is only a few atoms thick, were discovered more than 40 years ago [83, 84] when an external TM polarized plane wave was used to excite the channel. Field-effect transistor plasma waves have lately received great interest because of their application in engineering terahertz frequency sources and sensors [4, 137, 138, 139, 140]. Although research has mainly been focused in realizing far-infrared devices, nitride based heterostructures have been used for the mid-infrared frequencies [141]. More importantly, the frequency response of the device can be tuned by varying the gate voltage [142, 143].

In this paper, a nanoscale imaging technique is presented in which subwavelength plasma waves, generated by a current in a transistor channel that can be tuned by controlling gate voltage, are used as the illumination pattern required for SIM. The configuration effectively creates a much larger observable spatial frequency region as compared to a far-infrared (terahertz) plane wave. Due to the linear nature of the scheme, resolution of up to two orders of magnitude beyond the diffraction limit can be obtained with a weak field intensity. Although the wavelength is very small along the channel, it simultaneously brings the disadvantage of a fast decay in the vertical direction. The presence of a semiconductor layer above the channel further reduces the intensity of the wave. The resolution enhancement is ultimately limited by the layer thickness which should be kept as small as the fabrication process permits in order to reduce attenuation of the standing wave in the vertical direction.

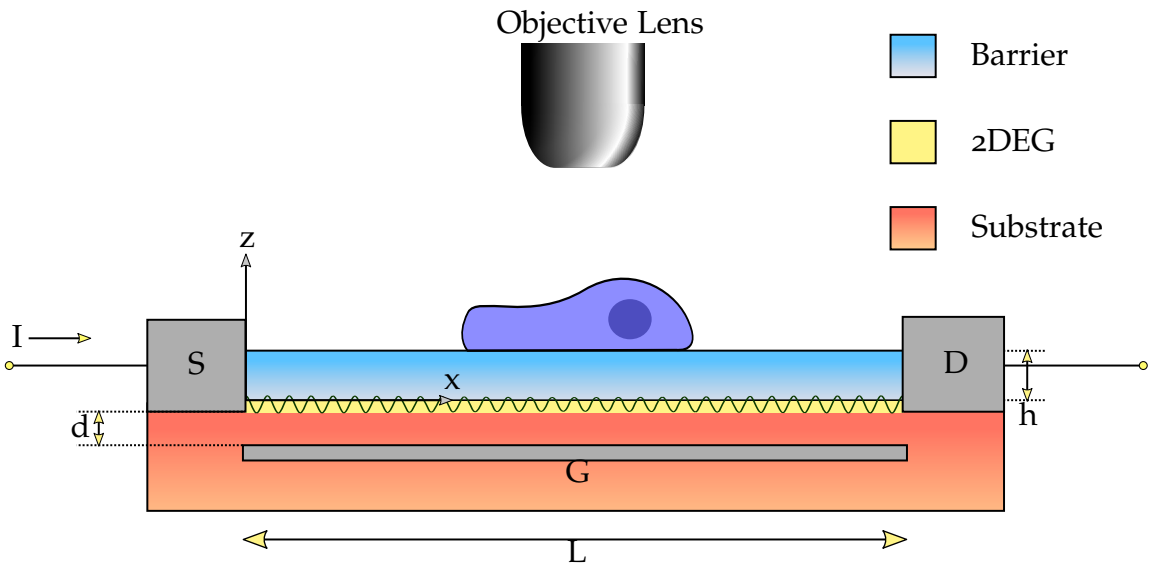


Figure 6.1: Illustration of the imaging scheme where sample is excited by plasmonic standing wave pattern generated by a current-driven instability in the channel

6.2 Theory

6.2.1 Dispersion Relation

A schematic diagram of the proposed system which is similar to a transistor, is shown in Fig. 6.1 where a 2DEG that acts as a transistor channel is formed at the interface of two semiconductor materials of slightly different band-gap energies. Plasma waves are generated in the channel when the source and drain terminals are driven by a current source. Due to reflections from the conducting boundaries, the channel region forms a cavity in which the plasma waves form a standing wave. The structure is backed by a gate terminal that spans the length L of the channel and spaced a distance d below the 2DEG. The gate capacitively couples with the 2DEG, and by varying the voltage, the velocity as well as concentration of electrons in the channel can be controlled. A barrier layer of thickness

h separates the sample from the 2DEG.

The compressed nature of the plasma waves can be described by its dispersion relation. Here, we consider the sample stage which is terminated by a gate at the bottom and free-space at the top, excited by a TM polarized plane wave. The effects of the drain and source terminals are ignored by assuming the layered semiconductor structure along with the 2DEG channel to be of infinite lateral extent. The dispersion relation which shows a frequency dependent resonance response of plasma waves in the 2D channel, is obtained by imposing the transverse resonance condition realized by an equivalent transmission line (TL) circuit [81, 2]. The 2DEG is modeled as a shunt admittance related to Drude-type surface conductivity [100],

$$Y_\sigma = \sigma_s = \frac{N_s e^2 \tau}{m^*} \frac{1}{1 + j\omega\tau}, \tag{6.1}$$

where N_s is the surface electron density in the channel, e is the electron charge, m^* is the effective electron mass in the heterostructure, τ is the scattering time of electrons, and ω is the angular frequency. The dispersion relation is then written as [105]:

$$Y^{\uparrow}(z_0) + Y^{\downarrow}(z_0) + Y_\sigma = 0. \tag{6.2}$$

Here, $Y^{\uparrow}(z_0)$ and $Y^{\downarrow}(z_0)$ are the up- and down-looking TL admittances from the 2DEG located at $z = 0$, and expressed as:

$$Y^{\uparrow}(z_0) = Y_2 \frac{1 - \Gamma^{\uparrow}(z_0)}{1 + \Gamma^{\uparrow}(z_0)}, \tag{6.3a}$$

$$Y^{\downarrow}(z_0) = -jY_1 \cot(k_{z1} h). \tag{6.3b}$$

For each layer, Y_i and k_{zi} where $i = 0, 1, 2$ are the respective TM mode admittance and transverse wavenumber of free-space, barrier and substrate layers respec-

tively given by:

$$Y_i = \frac{\omega \varepsilon_i \varepsilon_0}{k_{zi}}, \qquad k_{zi} = \pm\sqrt{k_0^2 \varepsilon_i - k_x^2} \tag{6.4}$$

where ε_i is the relative permittivity of i^{th} layer and k_x is the longitudinal propagation constant of the structure. The upward-looking reflection coefficient $\Gamma^{\uparrow}$ in (6.3a) is expressed in terms of the TM mode admittances:

$$\Gamma^{\uparrow}(z_0) = \frac{Y_1 - Y_0}{Y_0 + Y_1} \mathrm{e}^{-2jk_{z2}d_1} \tag{6.5}$$

A closed-form expression for the longitudinal propagation constant k_x determined by solving (5.2) is tedious, therefore numerical root-finding techniques such as the Newton method [114] have to be employed. As an example, the dispersion diagram of a AlGaN/GaN heterostructure is shown in Fig. 2.6 with no gate bias applied along with a light line whose slope is multiplied by a factor of 90 for illustration. The real part corresponds to propagation and imaginary part accounts for decay of the wave. The result is similar to the dispersion curve of gated 2DEG plasmons in which the gate terminal is located at the top [144, 145]. At 25 THz, the plasmon wavenumber is approximately 80 times greater than the free-space propagation constant for a plane wave. Furthermore, the loss which is related to the imaginary part of the plasmon wavenumber is also negligible. A sample observed through such a highly subwavelength illumination makes super-resolution possible. In Fig. 6.2b, control of the resonant frequency by varying the gate voltage is shown, using (6.6) and (6.7). Compared to a dipole in free-space, reducing the channel length increases the resonant frequency. The structure parameters used to compute the dispersion relation via (6.1)-(6.5) are

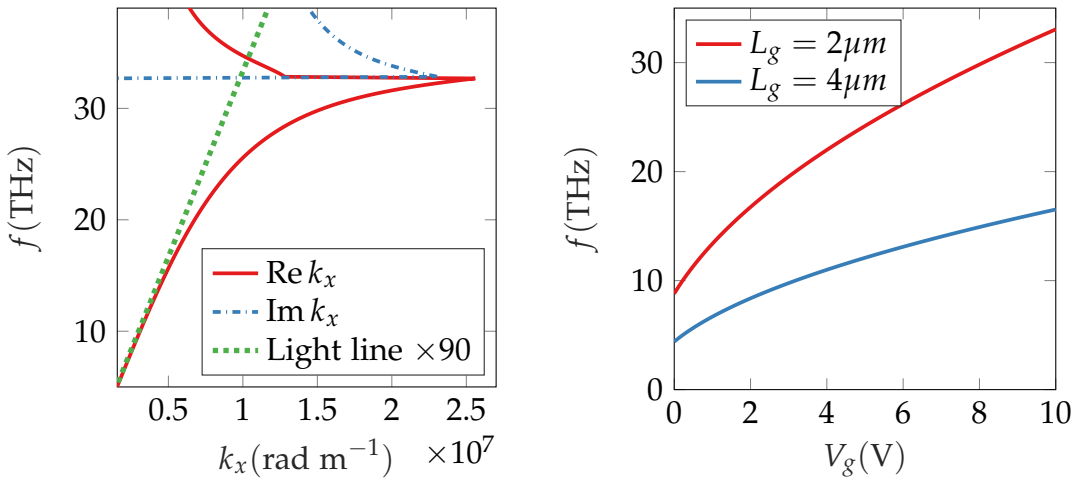


Figure 6.2: (a) Plasma wave dispersion diagram for a transistor structure supporting a 2DEG channel. (b) Effect of gate voltage on resonant frequency

now briefly discussed.

The gate-channel separation is $d = 100\,\text{nm}$. The channel length L is $2\,\mu\text{m}$ whereas the AlGaN barrier layer is $h = 20\,\text{nm}$ wide. The permittivity of both semiconductor layers is approximated to the static value, i.e., $\varepsilon_1 \approx \varepsilon_2 = 9.5$. Here the mole-fraction of aluminum in AlGaN alloy is 0.2000 [92]. A surface carrier density of $N_s = 5 \times 10^{13}\,\text{cm}^{-2}$ and scattering time τ of 114 ps corresponding to a temperature of 3 K is assumed. As the temperature is increased, τ gets smaller which leads to reduced mobility and introduces loss in the channel. Through the gate voltage V_g, the electron density N_s of the channel can be varied using:

$$N_s = N_0 \times \left(1 - \frac{V_g}{V_T}\right), \tag{6.6}$$

where N_0 is the zero-bias density and V_T is the gate threshold voltage of the transistor. For a channel terminated by highly conducting source and drain terminals

at each side, the resonant frequency as a function of carrier density is expressed as [146]:

$$\omega = \sqrt{\frac{N_s e^2 d}{m_* \varepsilon}} \frac{\pi}{L} \tag{6.7}$$

where ε is the average permittivity of the surrounding media. The tunability of plasma waves obtained by using (6.6) and (6.7), and assuming a gate threshold voltage of $-0.7640\,\mathrm{V}$ is shown in Fig. 6.2b. As expected, increasing the length of the channel reduces the resonant frequency.

6.2.2 Image Reconstruction

As seen in Fig. 6.5a, the plasma wave illumination pattern $I(\mathbf{r})$ can be assumed sinusoidal and expressed as:

$$I(\mathbf{r}) = 1 + \cos\Big(\mathbf{k}_\rho \cdot \mathbf{r} + \phi\Big) \tag{6.8}$$

where $\mathbf{k}_\rho = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}}$ is the spatial frequency wavevector, $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}$ is the two-dimensional positional vector and ϕ is the pattern phase. An image $M(\mathbf{r})$ of a sample atom distribution $F(\mathbf{r})$ observed through a microscope can be expressed as:

$$M(\mathbf{r}) = \big[F(\mathbf{r}) \cdot I(\mathbf{r})\big] \otimes H(\mathbf{r}) \tag{6.9}$$

where $H(\mathbf{r})$ is the point spread function (PSF) of the microscope, and $\cdot, \otimes$ denote multiplication and convolution operations in the spatial domain respectively. A frequency domain representation of the image obtained by taking the Fourier transform of (6.9) is expressed as:

$$\begin{aligned} \tilde{M}(\mathbf{k}) &= \Big[\tilde{F}(\mathbf{k}) \otimes \tilde{I}(\mathbf{k})\Big] \cdot \tilde{H}(\mathbf{k}) \\ &= \frac{1}{2}\Big[2\tilde{F}(\mathbf{k}) + \tilde{F}(\mathbf{k} - \mathbf{k}_\rho)e^{-j\phi} + \tilde{F}(\mathbf{k} + \mathbf{k}_\rho)e^{j\phi}\Big] \cdot \tilde{H}(\mathbf{k}) \end{aligned} \tag{6.10}$$

where $\sim$ over the letters indicates a frequency domain term and $\tilde{H}(k)$ is the optical transfer function (OTF) of the microscope. A spatial frequency representation of the scheme is illustrated in Fig. 6.3. In this scheme, we assume that the numerical aperture of the objective lens is unity, in which case the OTF is described by a circular disc as shown in Fig. 6.3(a) where the passband is bounded by $\sqrt{k_x^2 + k_y^2} = 2k_0 = \nu$. Since the plasmon frequency falls in the mid-infrared region, a relatively small k_0 implies that we need to image the sample a large number of times, which results in a slow imaging process. The process can be expedited by using an additional illumination such as a laser with frequency ω_v in the visible region. As shown in Fig. 6.4, the molecules in the sample are first excited from the ground state, $|g\rangle$ to the energy level $|e\rangle$ by using a laser of frequency ω_v. The plasmonic pattern then excites the molecules to an additional level, $|a\rangle$. Utilizing the spontaneous decay of the molecules from $|a\rangle$, we image the sample with photons of frequency $\omega_{ac} = \omega_a - \omega_c$. Here, ω_a and ω_c are the respective frequencies of energy levels $|a\rangle$ and $|c\rangle$. A frequency-selective photonic crystal is placed behind the objective lens to filter the photons of different frequencies. As a consequence of the preceding discussion, the resulting passband in the spatial frequency is now bounded by $\sqrt{k_x^2 + k_y^2} = 2k_{ac} = 2\omega_{ac}/c = \kappa$, which is the larger circle illustrated in Fig. 6.3. Since the frequency κ is much larger than ω_v, high resolution can be realized by imaging the sample only a few number of times. As evident in (6.10), a sinusoidal illumination pattern has three frequency components, which generate an image that is a linear combination of the sample along with two shifted versions as shown in Fig. 6.3b. To reconstruct the sample, three different images need to be captured with each possessing a different phase term ϕ. The process can be expressed as a system of linear equa-

tions,

$$\tilde{H}(\mathbf{k}) \cdot \begin{bmatrix} \tilde{F}(\mathbf{k}) \\ \tilde{F}(\mathbf{k}-\mathbf{k}_\rho) \\ \tilde{F}(\mathbf{k}+\mathbf{k}_\rho) \end{bmatrix} = \begin{bmatrix} 2 & e^{-j\phi_1} & e^{j\phi_1} \\ 2 & e^{-j\phi_2} & e^{j\phi_2} \\ 2 & e^{-j\phi_3} & e^{j\phi_3} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{M}_1(\mathbf{k}) \\ \tilde{M}_2(\mathbf{k}) \\ \tilde{M}_3(\mathbf{k}) \end{bmatrix} \tag{6.11}$$

The phase shifts in (6.11) are known beforehand. Frequency content of the sample up to k_ρ can therefore be observed due the Moiré effect which transports the high frequency information in to the observation region. To achieve two-dimensional enhancement in resolution, either the sample must be rotated about the optical axis of the microscope or the angular distribution of the illumination needs to be varied. In order to solve the three components of the spatial frequency as shown in above equation, we need to shift the plasmonic patterns. Our simulation results show that an additional incident plane wave can shift the pattern efficiently and the amount of shift is controlled by varying the angle of incidence with respect to the optical axis. Figure 6.5a shows a plot of normalized intensities. The first curve is the standing wave pattern obtained at the top surface which is only due to the surface current in the channel below. The remaining curves show the shifting of the standing wave, achieved by an additional plane wave incident from the top of the structure, whose intensity is greater than that of the plasma wave in the channel. By changing the angle of incidence slightly, different phase variations are acquired that are needed to solve the spatial frequency in (6.11). Only a small portion of the channel is illustrated in Figs. 6.5a and 6.5b, because the plasmonic wavelength is much smaller than the channel length L. A full-wave simulator (COMSOL) was used to plot the results. In this scheme, an external plane wave is used to shift the plasma wave pattern laterally in the sample stage. The electric field of a TM polarized plane wave is expressed

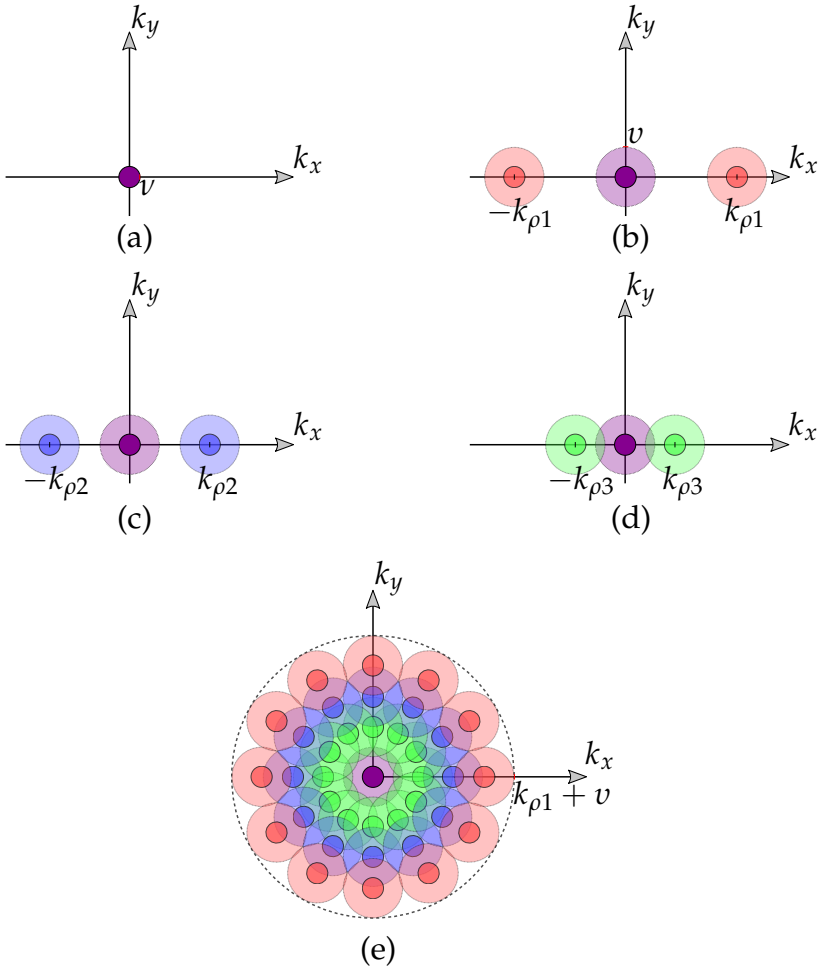


Figure 6.3: Resolution enhancement through SIM: (a) Diffraction limited observable region in frequency domain. Moiré effect using a sinusoidal illumination pattern bringing high frequency content under the observable region. (b)-(d) Sample illuminated at different plasma frequencies. (e) Effective resolution enhancement of $k_{\rho 1} + \kappa$ in two dimensions can be obtained after rotating the sample with respect to the optical axis

as, $\mathbf{E}_{ext} = \hat{\mathbf{x}}\, a + \hat{\mathbf{z}}\, b$. The total field at the surface is the sum of the plasmonic field and external plane wave. The total intensity is then expressed as:

$$\begin{aligned} |E|^2 &= (a + \cos k_\rho x)^2 + (b + \sin k_\rho x)^2 \\ &= a^2 + b^2 + 1 + 2\chi \cos\Big(k_\rho x + \psi\Big) \end{aligned} \tag{6.12}$$

where k_ρ is the spatial frequency of plasma wave, $\chi = \sqrt{a^2 + b^2}$ and $\psi = \text{atan}(b/a)$. The field components a and b are controlled by changing the incident angle of the plane wave. Using a commercial full-wave electromagnetic simulation tool (COMSOL Multiphysics), lateral shifting of the pattern is shown in Fig. 6.5a. Since the wavelength is much smaller than the channel length, only a small portion is shown. is As discussed earlier, the standing wave pattern can be tuned to different frequencies through gate voltage control.

Simulated standing wave patterns tuned to different frequencies by varying the gate voltage are shown in Fig. 6.5b. The resulting change in electron density modifies the surface conductivity (6.1) of the 2DEG and the dielectric function $\varepsilon(\omega) = 1 - j\sigma_s/(\omega\Delta\varepsilon_r)$ [122] where Δ is the 2DEG thickness and ε_r is the permittivity of the surrounding media.

The 2DEG channel essentially behaves as a cavity due to the resonance effects introduced by the two conducting boundaries, i.e., drain and source terminals. Therefore, the plasmon wavenumber, k_ρ can not be varied in a continuous fashion to cover all spatial frequencies. As an example, we set the length of the heterostructure and the resulting 2DEG channel to 2 µm in the simulation. Like any resonating structure, the plasmonic modes of the 2DEG channel are well-defined

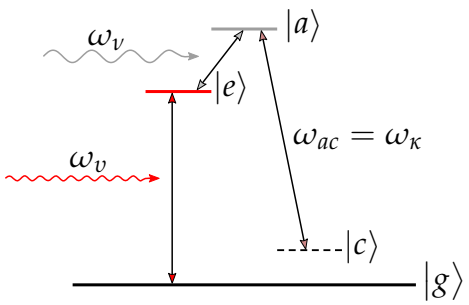
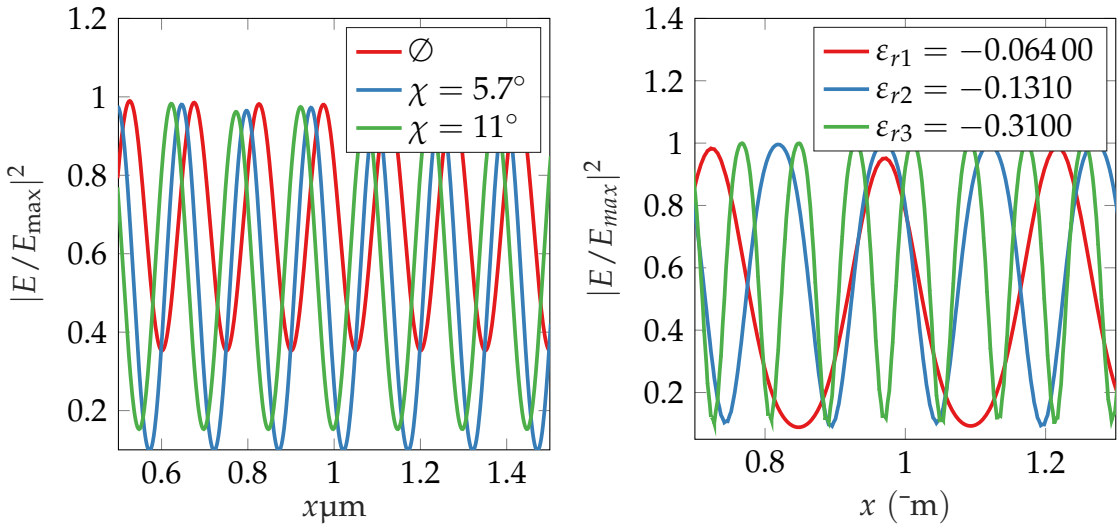


Figure 6.4: Sample excited to additional energy levels to enhance the spontaneous decay, ω_κ.



(a) (b)

Figure 6.5: Full-wave simulation results: (a) Phase shift achieved by exciting the structure with an additional illumination at an angle. (b) Tuning of the standing wave structure by applying gate bias

[147, 146, 92]. In terms of the wavenumber, the mode separation is approximately, $\Delta_k = 2\pi/(2 \times 10^{-6})$ rad/m. Thus, full coverage of all the spatial frequencies, just by tuning the plasma wave, cannot be accomplished. However, as discussed earlier, through the larger circle shown in Fig. 6.3, all the spatial frequencies up to the plasmon wavenumber can be recovered. For a laser having a wavelength of 600 nm, the circle radius is $\kappa = 4\pi/(6 \times 10^{-7})$ rad/m. We note that $\Delta_k \ll \kappa$, meaning that fewer images of the sample are required to realize super-resolution.

6.3 Simulated Results

We consider a sample with atom distribution shown in Fig. 6.6a. Each particle shown has a diameter of 1 nm. The minimum separation between the atoms is 38 nm and maximum is 137 nm. The 2D plasma waves are generated by a

dc current-driven instability that can also be equivalently excited by a TM polarized plane wave having a frequency of 25 THz, $\hbar\omega = 0.1000\,\text{eV}$ at zero gate bias. It is assumed that the numerical aperture (NA) is 1. To demonstrate the super-resolution technique, the atom distribution is first Fourier transformed to the spatial frequency domain. The critical step to achieve super-resolution reconstruction involves computation of the inverse Fourier transform using frequency content up to a circular region of radius $2k_\rho + \kappa$, where k_ρ can be varied by gate voltage. In Figs. 6.6b and 6.6c, a resolution enhancement 39.5 and 80 is shown corresponding to 152 nm and 74.90 nm resolution respectively.

Figure. 6.6b shows that the particles that are separated by a distance less than the resolution can not be resolved and appear as a contiguous blurry streak, whereas they are distinguishable in Fig. 6.6c. The sample is imaged about 20 times, which in terms of imaging speed is very fast and can be compared with nonlinear SIM [129]. However, unlike non-linear SIM, we use a weak illumination intensity in our scheme. At cryogenic temperatures, the plasmons exhibit near loss-less behavior in the 2DEG where the real part of the wavenumber is at least three orders of magnitude greater than the imaginary part. However, as the temperature rises, the reduced electron mobility resulting due to electron scattering introduces a plasma wave decay factor such . To account for the loss, the plasma

In the previous simulations, we didn't consider the loss in the 2DEG channel, which can introduce irregularities in the field pattern. These irregularities result in an erroneous solution for the three frequency components in (6.11), which are the circular regions shown in Fig. 6.3b, subsequently inducing image distortion. However, with ever improving nanofabrication processing techniques, the loss along the 2DEG channel can be very small especially at low temperature,

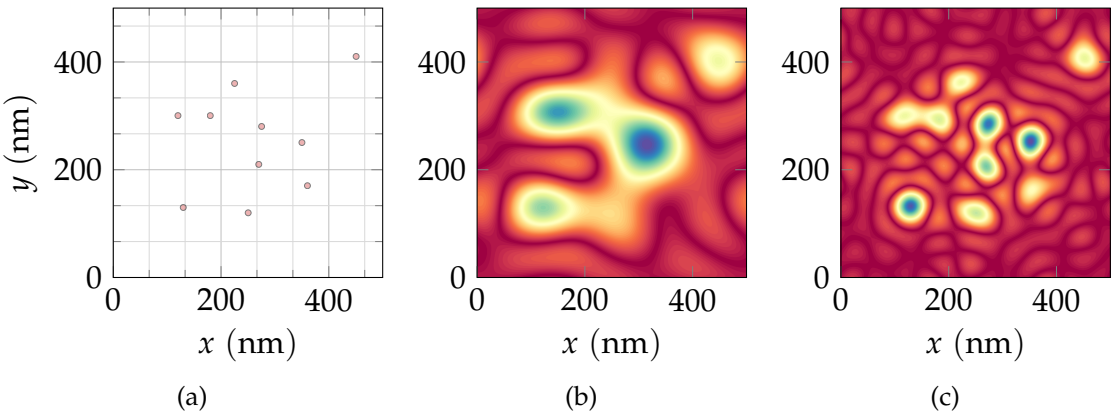


Figure 6.6: (a) Sample distribution. Simulation of the reconstructed sample image at: (b) $\mathrm{Re}\, k_\rho = 39.5$ (c) $\mathrm{Re}\, k_\rho = 80$

which subsequently means a large $\mathrm{Re}(k_\rho)/\,\mathrm{Im}(k_\rho)$ ratio. If the sample size is much smaller than the plasma wave propagation length, $1/\,\mathrm{Im}(k_\rho)$, the irregularities resulting from the loss can be neglected. We propose a super-resolution nanoscopy scheme based on the subwavelength surface electromagnetic plasma found in a semiconductor heterostructure. This method is useful in particular for light-sensitive samples as it requires a weak field intensity for illumination. In comparison with the metal based SIM where the plasmonic pattern is fixed by a determined structure, we can control the period of the plasmonic pattern by varying the gate voltage. Moreover, in contrast to graphene based SIM, our scheme uses surface current to excite the plasmons, which does not require any wavenumber matching mechanism like gratings. The image reconstruction algorithm is as efficient as in conventional linear SIM, yet the high wavenumber along with the tunability of the plasma waves allows it to achieve subdiffraction resolution of up to 80 times in the mid-infrared range.

7. CONCLUSION AND FUTURE RECOMMENDATIONS

In this work, plasmonic structures with an emphasis on antenna designs and imaging systems were studied that support subwavelength wave phenomena in the optical as well as the terahertz frequency domain. It was shown that the plasmonic antenna designs discussed herein provide an ideal outlet for realization of miniaturized communication devices. An in-depth theory dealing with the wave propagation mechanism in the form of surface plasmons was presented in the optical frequency domain. It was identified that the noble metals exhibit an optical frequency dielectric function with a negative real part, which is a necessary condition for the existence of surface plasmons. Likewise in the terahertz frequency region, plasmonic activity was observed in the few atoms wide electron channel of a high electron mobility transistor which is composed of an epitaxially grown semiconductor heterostructure.

A full wave analysis of plasmonic structures, and in particular semiconductor heterostructures, is inefficient using commercial software that are mostly based on differential equation discretization such as FEM and FDTD due to the presence of an extremely thin conducting layer. On the other hand, integral equation techniques in which the fields are computed by first constructing an appropriate Green function that is representative of the physical structure, and then followed by a method of moments discretization, are computationally efficient and most importantly, provide a great deal of insight on the wave mechanism in the structure. It must be stated that the mathematical formulation of integral equations is more involved and the integration routine must be cognizant of the singularities present in the integration kernels. In this work, an equivalent transmission

line approach was followed to formulate the Green functions of an infinitesimally thin conductive sheet embedded in a semiconductor heterostructure. The complex plane integration was performed along the positive real line where the branch point singularities were circumvented by a triangular deformation of the integration path. Moreover, the mixed potentials formulation was adopted owing to comparatively weaker singular nature of the integral kernels. The results obtained by computing the Sommerfeld integrals extracted from a magnetic vector potential formulation for a free-standing conductive sheet show the existence of surface plasmons in the terahertz frequency regime.

The subwavelength nature of plasmonic structures was underlined by the dispersion relations and the resultant dispersion curves. Surface plasmons existing at a metal-dielectric interface at optical frequencies yield an analytical solutions of the dispersion relation. The dispersion relations for more complex semiconductor heterostructures were numerically solved in the terahertz frequency region using a robust complex-valued root-finding technique called the argument principle method. The results showed a much higher confinement of plasmons in the semiconductor heterostructure than the metal-dielectric interface, mainly due to the two-dimensional wave nature in the former case. In this regard, a super-resolution imaging scheme using a periodic structured illumination was proposed. It was shown that the terahertz standing plasma waves generated along the heterointerface, laterally enclosed by the transistor, can be used to resolve a sample with particle separation in the range of a few nanometers.

7.1 Recommendations for Future Work

Terahertz plasmonics is currently being seen as the brightest prospect in terms of creating efficient terahertz devices that include sources and sensors. The

thin plasma region inside a semiconductor heterostructure exhibits a resonant response in the terahertz frequency region that can be tuned using an active transistor environment. Unfortunately, at present practical efficiencies in terms of power can only be obtained at very low temperatures. Currently, most of the heterostructures are made from group III-V materials and their associated alloys. Extensive research is on-going to explored heterostructures that can operate at higher temperature. In this regard, materials such as perovskites and dichalcogenides have lately received an increased level of interest. Furthermore, the two-dimensional nature of the plasmonic structure is slowly evolving into a new research field termed as *metasurfaces* [148, 149].

Various aspects of the analysis and design methods described throughout this dissertation be greatly improved. The surface conductivity of the 2DEG was assumed to be a scalar quantity. The quantum effects associated with a 2DEG due to external electric or magnetic fields can be incorporated in to the surface conductivity by modeling it as a tensor quantity.

The root-finding technique discussed in Section 4 currently involves a considerable amount of guess work in determining whether the poles are proper or improper. The routine can be improved by mapping the complex plane into a new coordinate system using a trigonometric transformation that essentially removes the branch points and its associated branch cuts.

REFERENCES

[1] P. B. Johnson and R. W. Christy, "Optical constants of the noble metals," *Physical Review B*, vol. 6, no. 12, pp. 4370–4379, Dec. 1972.

[2] K. A. Michalski, "Electromagnetic field computation in planar multilayers," in *Encyclopedia of RF and Microwave Engineering*. John Wiley & Sons, Inc., Apr. 2005.

[3] K. Kempa, P. Bakshi, J. Cen, and H. Xie, "Spontaneous generation of plasmons by ballistic electrons," *Physical Review B*, vol. 43, no. 11, pp. 9273–9274, Apr. 1991.

[4] M. Dyakonov and M. Shur, "Shallow water analogy for a ballistic field effect transistor: New mechanism of plasma wave generation by dc current," *Physical Review Letters*, vol. 71, no. 15, pp. 2465–2468, Oct. 1993.

[5] M. Dyakonov and M. S. Shur, "Plasma wave electronics for terahertz applications," in *Terahertz Sources and Systems*. Springer Nature, 2001, pp. 187–207.

[6] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno, "Terahertz quantum-cascade laser at $\lambda \approx 100\mu$m using metal waveguide for mode confinement," *Applied Physics Letters*, vol. 83, no. 11, pp. 2124–2126, Sep. 2003.

[7] M. Krasheninnikov and A. Chaplik, "Instabilities of two-dimensional plasma waves," *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki*, vol. 79, pp. 555–560, Aug. 1980.

[8] K. Michalski and J. Mosig, "Multilayered media Green's functions in integral equation formulations," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 3, pp. 508–519, Mar. 1997.

[9] F. Ouyang, P. E. Batson, and M. Isaacson, "Quantum size effects in the surface-plasmon excitation of small metallic particles by electron-energy-loss spectroscopy," *Physical Review B*, vol. 46, no. 23, pp. 15 421–15 425, Dec. 1992.

[10] S. Nie, "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering," *Science*, vol. 275, no. 5303, pp. 1102–1106, Feb. 1997.

[11] K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M. S. Feld, "Single molecule detection using surface-enhanced Raman scattering (SERS)," *Physical Review Letters*, vol. 78, no. 9, pp. 1667–1670, Mar. 1997.

[12] F. Arduini, A. Amine, D. Moscone, and G. Palleschi, "Biosensors based on cholinesterase inhibition for insecticides, nerve agents and aflatoxin B1 detection (review)," *Microchimica Acta*, vol. 170, no. 3-4, pp. 193–214, Mar. 2010.

[13] R. Nevels, G. Welch, P. Cremer, P. Hemmer, T. Phillips, S. Scully, A. Sokolov, A. Svidzinsky, H. Xia, A. Zheltikov, and M. Scully, "Figuration and detection of single molecules," *Molecular Physics*, vol. 110, no. 15-16, pp. 1993–2000, Aug. 2012.

[14] R. Adato, A. A. Yanik, and H. Altug, "On chip plasmonic monopole nano-antennas and circuits," *Nano Letters*, vol. 11, no. 12, pp. 5219–5226, Dec. 2011.

[15] C. S. Torres, S. Zankovych, J. Seekamp, A. Kam, C. C. Cedeño, T. Hoffmann, J. Ahopelto, F. Reuther, K. Pfeiffer, G. Bleidiessel, G. Gruetzner, M. Maximov, and B. Heidari, "Nanoimprint lithography: an alternative nanofabrication approach," *Materials Science and Engineering: C*, vol. 23, no. 1-2, pp. 23–31, Jan. 2003.

[16] K. Ishihara, K. Ohashi, T. Ikari, H. Minamide, H. Yokoyama, J. Ichi Shikata, and H. Ito, "Terahertz-wave near-field imaging with subwavelength resolution using surface-wave-assisted bow-tie aperture," *Applied Physics Letters*, vol. 89, no. 20, p. 201120, Nov. 2006.

[17] R. D. Nevels and K. A. Michalski, "On the behavior of surface plasmons at a metallo-dielectric interface," *Journal of Lightwave Technology*, vol. 32, no. 19, pp. 3299–3305, Oct. 2014.

[18] J. N. Farahani, "Single emitters coupled to bow-tie nano-antennas," Ph.D. dissertation, Institut für Physik Universität Basel, 2006.

[19] M. J. Weber, *Handbook of Lasers (Laser & Optical Science & Technology)*. CRC Press, 2000.

[20] R. H. Ritchie, "Plasma losses by fast electrons in thin films," *Physical Review*, vol. 106, no. 5, pp. 874–881, Jun. 1957.

[21] A. Otto, "Spectroscopy of surface polaritons by attenuated total reflection," *Optical Properties of Solids, New Developments*, pp. 677–727, 1976.

[22] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer Berlin Heidelberg, 1988.

[23] K. L. Kelly, E. Coronado, L. L. Zhao, and G. C. Schatz, "The optical properties of metal nanoparticles: the influence of size, shape, and dielectric

environment," *The Journal of Physical Chemistry B*, vol. 107, no. 3, pp. 668–677, Jan. 2003.

[24] S. A. Maier and H. A. Atwater, "Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures," *Journal of Applied Physics*, vol. 98, no. 1, p. 011101, Jul. 2005.

[25] Q.-H. Park, "Optical antennas and plasmonics," *Contemporary Physics*, vol. 50, no. 2, pp. 407–423, Mar. 2009.

[26] J. Homola, Ed., *Surface Plasmon Resonance Based Sensors*. Springer Berlin Heidelberg, 2006.

[27] E. W. Max Born, *Principles of Optics*. Cambridge University Pr., 2002.

[28] A. Moroz, "Non-radiative decay of a dipole emitter close to a metallic nanoparticle: Importance of higher-order multipole contributions," *Optics Communications*, vol. 283, no. 10, pp. 2277–2287, May 2010.

[29] C. F. Bohren and D. R. Huffman, Eds., *Absorption and Scattering of Light by Small Particles*. Wiley-VCH Verlag GmbH, Apr. 1998.

[30] A. Archambault, T. V. Teperik, F. Marquier, and J. J. Greffet, "Surface plasmon Fourier optics," *Physical Review B*, vol. 79, no. 19, May 2009.

[31] D. W. Lynch and W. Hunter, "Comments on the optical constants of metals and an introduction to the data for several metals," in *Handbook of Optical Constants of Solids*. Elsevier, 1997, pp. 275–367.

[32] K. A. Michalski, "On the low-order partial-fraction fitting of dielectric functions at optical wavelengths," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 12, pp. 6128–6135, Dec. 2013.

[33] M. Durach, A. Rusina, M. I. Stockman, and K. Nelson, "Toward full spatiotemporal control on the nanoscale," *Nano Letters*, vol. 7, no. 10, pp. 3145–3149, Oct. 2007.

[34] J. J. Burke, G. I. Stegeman, and T. Tamir, "Surface-polariton-like waves guided by thin, lossy metal films," *Physical Review B*, vol. 33, no. 8, pp. 5186–5201, Apr. 1986.

[35] D. Sarid, "Long-range surface-plasma waves on very thin metal films," *Physical Review Letters*, vol. 47, no. 26, pp. 1927–1930, Dec. 1981.

[36] P. Berini, "Plasmon-polariton waves guided by thin lossy metal films of finite width: Bound modes of symmetric structures," *Physical Review B*, vol. 61, no. 15, pp. 10 484–10 503, Apr. 2000.

[37] ——, "Plasmon-polariton waves guided by thin lossy metal films of finite width: Bound modes of asymmetric structures," *Physical Review B*, vol. 63, no. 12, Mar. 2001.

[38] F. Yang, J. R. Sambles, and G. W. Bradberry, "Long-range surface modes supported by thin films," *Physical Review B*, vol. 44, no. 11, pp. 5855–5872, Sep. 1991.

[39] T. H. Taminiau, R. J. Moerland, F. B. Segerink, L. Kuipers, and N. F. van Hulst, "$\lambda/4$ resonance of an optical monopole antenna probed by single molecule fluorescence," *Nano Letters*, vol. 7, no. 1, pp. 28–33, Jan. 2007.

[40] L. Novotny, "Effective wavelength scaling for optical antennas," *Physical Review Letters*, vol. 98, no. 26, Jun. 2007.

[41] T. Søndergaard and S. Bozhevolnyi, "Slow-plasmon resonant nanostructures: Scattering and field enhancements," *Physical Review B*, vol. 75, no. 7,

Feb. 2007.

[42] T. Søndergaard, J. Beermann, A. Boltasseva, and S. I. Bozhevolnyi, "Slow-plasmon resonant-nanostrip antennas: Analysis and demonstration," *Physical Review B*, vol. 77, no. 11, Mar. 2008.

[43] N. Engheta, A. Salandrino, and A. Alù, "Circuit elements at optical frequencies: Nanoinductors, nanocapacitors, and nanoresistors," *Physical Review Letters*, vol. 95, no. 9, Aug. 2005.

[44] A. Alù, A. Salandrino, and N. Engheta, "Coupling of optical lumped nanocircuit elements and effects of substrates," *Optics Express*, vol. 15, no. 21, p. 13865, 2007.

[45] Y. Zhao, N. Engheta, and A. Alù, "Effects of shape and loading of optical nanoantennas on their sensitivity and radiation properties," *Journal of the Optical Society of America B*, vol. 28, no. 5, p. 1266, Apr. 2011.

[46] M. Agio and A. Alù, Eds., *Optical Antennas*. Cambridge University Press, 2013.

[47] P. J. Schuck, D. P. Fromm, A. Sundaramurthy, G. S. Kino, and W. E. Moerner, "Improving the mismatch between light and nanoscale objects with gold bowtie nanoantennas," *Physical Review Letters*, vol. 94, no. 1, Jan. 2005.

[48] H. Fischer and O. J. F. Martin, "Engineering the optical response of plasmonic nanoantennas," *Optics Express*, vol. 16, no. 12, p. 9144, Jun. 2008.

[49] O. L. Muskens, V. Giannini, J. A. Sánchez-Gil, and J. G. Rivas, "Optical scattering resonances of single and coupled dimer plasmonic nanoantennas," *Optics Express*, vol. 15, no. 26, p. 17736, 2007.

[50] H. F. Hofmann, T. Kosako, and Y. Kadoya, "Design parameters for a nano-optical Yagi–Uda antenna," *New Journal of Physics*, vol. 9, no. 7, pp. 217–217, Jul. 2007.

[51] T. Kosako, Y. Kadoya, and H. F. Hofmann, "Directional control of light by a nano-optical Yagi–Uda antenna," *Nature Photonics*, vol. 4, no. 5, pp. 312–315, Mar. 2010.

[52] J. Dorfmüller, D. Dregely, M. Esslinger, W. Khunsin, R. Vogelgesang, K. Kern, and H. Giessen, "Near-field dynamics of optical Yagi–Uda nanoantennas," *Nano Letters*, vol. 11, no. 7, pp. 2819–2824, Jul. 2011.

[53] C. A. Balanis, *Antenna Theory*. John Wiley & Sons, 2016.

[54] R. S. Pavlov, A. G. Curto, and N. F. van Hulst, "Log-periodic optical antennas with broadband directivity," *Optics Communications*, vol. 285, no. 16, pp. 3334–3340, Jul. 2012.

[55] E. M. Purcell, H. C. Torrey, and R. V. Pound, "Resonance absorption by nuclear magnetic moments in a solid," *Physical Review*, vol. 69, no. 1-2, pp. 37–38, Jan. 1946.

[56] K. J. Vahala, "Optical microcavities," *Nature*, vol. 424, no. 6950, pp. 839–846, Aug. 2003.

[57] B.-S. Song, S. Noda, T. Asano, and Y. Akahane, "Ultra-high-Q photonic double-heterostructure nanocavity," *Nature Materials*, vol. 4, no. 3, pp. 207–210, Feb. 2005.

[58] T. Gaebel, I. Popa, A. Gruber, M. Domhan, F. Jelezko, and J. Wrachtrup, "Stable single-photon source in the near infrared," *New Journal of Physics*, vol. 6, pp. 98–98, Jul. 2004.

[59] A. G. Curto, "Optical antennas control light emission," Ph.D. dissertation, ICFO – Institut de Ciències Fotòniques Universitat Politècnica de Catalunya, 2013.

[60] S. A. Maier, "Plasmonic field enhancement and SERS in the effective mode volume picture," *Optics Express*, vol. 14, no. 5, p. 1957, 2006.

[61] J. Barthes, G. C. des Francs, A. Bouhelier, J.-C. Weeber, and A. Dereux, "Purcell factor for a point-like dipolar emitter coupled to a two-dimensional plasmonic waveguide," *Physical Review B*, vol. 84, no. 7, Aug. 2011.

[62] U. C. Fischer, "Submicrometer aperture in a thin metal film as a probe of its microenvironment through enhanced light scattering and fluorescence," *Journal of the Optical Society of America B*, vol. 3, no. 10, p. 1239, Oct. 1986.

[63] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays," *Nature*, vol. 391, no. 6668, pp. 667–669, Feb. 1998.

[64] H. A. Bethe, "Theory of diffraction by small holes," *Physical Review*, vol. 66, no. 7-8, pp. 163–182, Oct. 1944.

[65] C. J. Bouwkamp, "On Bethe's theory of diffraction by small holes," *Philips Journal of Research*, vol. 5, pp. 321–332, 1950.

[66] H. J. Lezec and T. Thio, "Diffracted evanescent wave model for enhanced and suppressed optical transmission through subwavelength hole arrays," *Optics Express*, vol. 12, no. 16, p. 3629, 2004.

[67] G. Gay, O. Alloschery, B. V. de Lesegno, C. O'Dwyer, J. Weiner, and H. J. Lezec, "The optical response of nanostructured surfaces and the composite

diffracted evanescent wave model," *Nature Physics*, vol. 2, no. 4, pp. 262–267, Mar. 2006.

[68] P. Lalanne and J. P. Hugonin, "Interaction between optical nano-objects at metallo-dielectric interfaces," *Nature Physics*, vol. 2, no. 8, pp. 551–556, Jul. 2006.

[69] R. Collin, "Hertzian dipole radating over a lossy earth or sea: some early and late 20^{th}-century controversies," *IEEE Antennas and Propagation Magazine*, vol. 46, no. 2, pp. 64–79, Apr. 2004.

[70] F. J. Garcia-Vidal, L. Martin-Moreno, T. W. Ebbesen, and L. Kuipers, "Light passing through subwavelength apertures," *Reviews of Modern Physics*, vol. 82, no. 1, pp. 729–787, Mar. 2010.

[71] L. Wang, S. M. Uppuluri, E. X. Jin, and X. Xu, "Nanolithography using high transmission nanoscale bowtie apertures," *Nano Letters*, vol. 6, no. 3, pp. 361–364, Mar. 2006.

[72] T. Grosjean, M. Mivelle, F. I. Baida, G. W. Burr, and U. C. Fischer, "Diabolo nanoantenna for enhancing and confining the magnetic optical field," *Nano Letters*, vol. 11, no. 3, pp. 1009–1013, Mar. 2011.

[73] J.-H. Kang, K. Kim, H.-S. Ee, Y.-H. Lee, T.-Y. Yoon, M.-K. Seo, and H.-G. Park, "Low-power nano-optical vortex trapping via plasmonic diabolo nanoantennas," *Nature Communications*, vol. 2, p. 582, Dec. 2011.

[74] O. Mahboub, S. C. Palacios, C. Genet, F. J. Garcia-Vidal, S. G. Rodrigo, L. Martin-Moreno, and T. W. Ebbesen, "Optimization of bull's eye structures for transmission enhancement," *Optics Express*, vol. 18, no. 11, p. 11292, May 2010.

[75] A. G. Curto, G. Volpe, T. H. Taminiau, M. P. Kreuzer, R. Quidant, and N. F. van Hulst, "Unidirectional emission of a quantum dot coupled to a nanoantenna," *Science*, vol. 329, no. 5994, pp. 930–933, Aug. 2010.

[76] N. Félidj, J. Aubard, G. Lévi, J. R. Krenn, A. Hohenau, G. Schider, A. Leitner, and F. R. Aussenegg, "Optimized surface-enhanced Raman scattering on gold nanoparticle arrays," *Applied Physics Letters*, vol. 82, no. 18, pp. 3095–3097, May 2003.

[77] J. G. V. Bladel, *Electromagnetic Fields*, ser. IEEE Press Series on Electromagnetic Wave Theory. John Wiley & Sons, 2007.

[78] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves (ieee Press Series on Electromagnetic Wave Theory)*, ser. IEEE Press Series on Electromagnetic Wave Theory. Wiley-IEEE Press, 1994.

[79] D. G. Fang, *Antenna Theory and Microstrip Antennas*. CRC PR INC, 2010.

[80] T. Itoh, "Spectral Domain Immitance Approach for Dispersion Characteristics of Generalized Printed Transmission Lines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 28, no. 7, pp. 733–736, Jul. 1980.

[81] R. Kastner, E. Heyman, and A. Sabban, "Spectral domain iterative analysis of single- and double-layered microstrip antennas using the conjugate gradient algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 36, no. 9, pp. 1204–1212, Sep. 1988.

[82] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, ser. Applied Mathematics Series. National Bureau of Standards. New York: Dover, 1968.

[83] F. Stern, "Polarizability of a two-dimensional electron gas," *Physical Review Letters*, vol. 18, no. 14, pp. 546–548, Apr. 1967.

[84] S. J. Allen, D. C. Tsui, and R. A. Logan, "Observation of the two-dimensional plasmon in silicon inversion layers," *Physical Review Letters*, vol. 38, no. 17, pp. 980–983, Apr. 1977.

[85] J. R. Wait, "The electromagnetic fields of a horizontal dipole in the presence of a conducting half-space," *Canadian Journal of Physics*, vol. 39, no. 7, pp. 1017–1028, Jul. 1961.

[86] K. Michalski, "On the efficient evaluation of integral arising in the Sommerfeld halfspace problem," *IEE Proceedings H Microwaves, Antennas and Propagation*, vol. 132, no. 5, p. 312, 1985.

[87] H. Takahasi and M. Mori, "Double exponential formulas for numerical integration," *Publications of the Research Institute for Mathematical Sciences*, vol. 9, no. 3, pp. 721–741, 1974.

[88] R. Golubovic, A. G. Polimeridis, and J. R. Mosig, "Efficient algorithms for computing Sommerfeld integral tails," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 5, pp. 2409–2417, 2012.

[89] I. D. Koufogiannis, M. Mattes, and J. R. Mosig, "On the Development and Evaluation of Spatial-Domain Green's Functions for Multilayered Structures With Conductive Sheets," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 1, pp. 20–29, Jan. 2015.

[90] K. Michalski, "Extrapolation methods for Sommerfeld integral tails," *IEEE Transactions on Antennas and Propagation*, vol. 46, no. 10, pp. 1405–1418, 1998.

[91] K. A. Michalski and J. R. Mosig, "Efficient computation of Sommerfeld integral tails–methods and algorithms," *Journal of Electromagnetic Waves and Applications*, vol. 30, no. 3, pp. 281–317, 2016.

[92] A. V. Muravjov, D. B. Veksler, V. V. Popov, O. V. Polischuk, N. Pala, X. Hu, R. Gaska, H. Saxena, R. E. Peale, and M. S. Shur, "Temperature dependence of plasmonic terahertz absorption in grating-gate gallium-nitride transistor structures," *Applied Physics Letters*, vol. 96, no. 4, p. 042105, Jan. 2010.

[93] G. W. Hanson, "Dyadic Green's functions and guided surface waves for a surface conductivity model of graphene," *Journal of Applied Physics*, vol. 103, no. 6, p. 064302, Mar. 2008.

[94] G. Lovat, "Equivalent circuit for electromagnetic interaction and transmission through graphene sheets," *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 1, pp. 101–109, Feb. 2012.

[95] J. S. Gomez-Diaz, J. R. Mosig, and J. Perruisseau-Carrier, "Effect of spatial dispersion on surface waves propagating along graphene sheets," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 7, pp. 3589–3596, Jul. 2013.

[96] T. Senior and J. L. Volakis, *Approximate Boundary Conditions in Electromagnetics (IEEE Electromagnetic Waves Series)*. The Institution of Engineering and Technology, 1995.

[97] D. J. Hoppe, *Impedance Boundary Conditions In Electromagnetics*. CRC Press, 1995.

[98] G. A. T. Warren L. Stutzman, *Antenna Theory and Design*. John Wiley & Sons INC, 2012.

[99] R. F. Harrington, *Field Computation Moment Methods*. IEEE COMPUTER SOC PR, 1993.

[100] P. J. Burke, I. B. Spielman, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, "High frequency conductivity of the high-mobility two-dimensional electron gas," *Applied Physics Letters*, vol. 76, no. 6, pp. 745–747, Feb. 2000.

[101] G. Herranz, F. Sánchez, N. Dix, M. Scigaj, and J. Fontcuberta, "High mobility conduction at (110) and (111) $LaAlO_3/SrTiO_3$ interfaces," *Scientific Reports*, vol. 2, p. 758, Oct. 2012.

[102] T. Senior, "Backscattering from resistive strips," *IEEE Transactions on Antennas and Propagation*, vol. 27, no. 6, pp. 808–813, 1979.

[103] J. Richmond, "Scattering by a dielectric cylinder of arbitrary cross section shape," *IEEE Transactions on Antennas and Propagation*, vol. 13, no. 3, pp. 334–341, May 1965.

[104] T. B. A. Senior and J. L. Volakis, "Sheet simulation of a thin dielectric layer," *Radio Science*, vol. 22, no. 7, pp. 1261–1272, Dec. 1987.

[105] J. S. Gómez-Díaz and J. Perruisseau-Carrier, "Propagation of hybrid transverse magnetic-transverse electric plasmons on magnetically biased graphene sheets," *Journal of Applied Physics*, vol. 112, no. 12, p. 124906, Dec. 2012.

[106] L. M. Delves and J. N. Lyness, "A numerical method for locating the zeros of an analytic function," *Mathematics of Computation*, vol. 21, no. 100, pp. 543–543, 1967.

[107] M. Carpentier and A. D. Santos, "Solution of equations involving analytic functions," *Journal of Computational Physics*, vol. 45, no. 2, pp. 210–220, Feb.

1982.

[108] L. Botten, M. Craig, and R. McPhedran, "Complex zeros of analytic functions," *Computer Physics Communications*, vol. 29, no. 3, pp. 245–259, May 1983.

[109] P. Kravanja, M. V. Barel, O. Ragos, M. Vrahatis, and F. Zafiropoulos, "ZEAL: A mathematical software package for computing zeros of analytic functions," *Computer Physics Communications*, vol. 124, no. 2-3, pp. 212–232, Feb. 2000.

[110] M. Dellnitz, O. Schütze, and Q. Zheng, "Locating all the zeros of an analytic function in one complex variable," *Journal of Computational and Applied Mathematics*, vol. 138, no. 2, pp. 325–333, Jan. 2002.

[111] C. Gillan, A. Schuchinsky, and I. Spence, "Computing zeros of analytic functions in the complex plane without using derivatives," *Computer Physics Communications*, vol. 175, no. 4, pp. 304–313, Aug. 2006.

[112] P. Y. Chen and Y. Sivan, "Robust Location of Optical Fiber Modes via the Argument Principle Method," *Computer Physics Communications*, vol. 214, no. 2, pp. 105–116, May 2017.

[113] S. G. Krantz, *The Argument Principle*. Boston, MA: Birkhäuser Boston, 1999, pp. 69–78.

[114] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*. Cambridge University Pr., 2007.

[115] L. Shampine, "Vectorized adaptive quadrature in MATLAB," *Journal of Computational and Applied Mathematics*, vol. 211, no. 2, pp. 131–140, Feb. 2008.

[116] P. Kravanja, T. Sakurai, and M. V. Barel, "On locating clusters of zeros of analytic functions," *Bit Numerical Mathematics*, vol. 39, no. 4, pp. 646–682, Dec. 1999.

[117] J. E. Gentle, "Computation of eigenvectors and eigenvalues and the singular value decomposition," in *Statistics and Computing*. Springer New York, 1998, pp. 123–135.

[118] K. A. Michalski, "Automatic, robust and efficient pole location for planar, uniaxial multilayers," 2006.

[119] A. G. Polimeridis, T. V. Yioultsis, and T. D. Tsiboukis, "An efficient pole extraction technique for the computation of Green's functions in stratified media using a sine transformation," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 1, pp. 227–229, Jan. 2007.

[120] Ş. E. Kocabaş, G. Veronis, D. A. B. Miller, and S. Fan, "Modal analysis and coupling in metal-insulator-metal waveguides," *Physical Review B*, vol. 79, no. 3, Jan. 2009.

[121] O. Sydoruk, J. B. Wu, A. Mayorov, C. D. Wood, D. K. Mistry, and J. E. Cunningham, "Terahertz plasmons in coupled two-dimensional semiconductor resonators," *Physical Review B*, vol. 92, no. 19, Nov. 2015.

[122] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Reviews of Modern Physics*, vol. 54, no. 2, pp. 437–672, Apr. 1982.

[123] H. Yoon, K. Y. M. Yeung, P. Kim, and D. Ham, "Plasmonics with two-dimensional conductors," *Philosophical Transactions of the Royal Society A:*

Mathematical, Physical and Engineering Sciences, vol. 372, no. 2012, pp. 20 130 104–20 130 104, Feb. 2014.

[124] O. Sydoruk, K. Choonee, and G. C. Dyer, "Transmission and reflection of terahertz plasmons in two-dimensional plasmonic devices," *IEEE Transactions on Terahertz Science and Technology*, vol. 5, no. 3, pp. 486–496, May 2015.

[125] M. Minsky, "Memoir on inventing the confocal scanning microscope," *Scanning*, vol. 10, no. 4, pp. 128–138, 1988.

[126] M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy." *Journal of Microscopy*, vol. 198, no. 2, pp. 82–87, May 2000.

[127] R. Heintzmann and C. G. Cremer, "Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating," in *Optical Biopsies and Microscopic Techniques III*, I. J. Bigio, H. Schneckenburger, J. Slavik, K. Svanberg, and P. M. Viallet, Eds., vol. 3568. SPIE, Jan. 1999, pp. 185–196.

[128] R. Heintzmann and G. Ficz, "Breaking the resolution limit in light microscopy," *Briefings in Functional Genomics and Proteomics*, vol. 5, no. 4, pp. 289–301, May 2006.

[129] M. G. L. Gustafsson, "Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution," *Proceedings of the National Academy of Sciences*, vol. 102, no. 37, pp. 13 081–13 086, Sep. 2005.

[130] H. Nassenstein, "Superresolution by diffraction of subwaves," *Optics Communications*, vol. 2, no. 5, pp. 231–234, Oct. 1970.

[131] F. Wei and Z. Liu, "Plasmonic structured illumination microscopy," *Nano Letters*, vol. 10, no. 7, pp. 2531–2536, Jul. 2010.

[132] X. Zeng, M. Al-Amri, and M. S. Zubairy, "Nanometer-scale microscopy via graphene plasmons," *Physical Review B*, vol. 90, no. 23, pp. 2–6, Dec. 2014.

[133] X. Zeng, L. Fan, and M. S. Zubairy, "Deep-subwavelength lithography via graphene plasmons," *Phys. Rev. A*, vol. 95, no. 5, May 2017.

[134] A. B. Mikhailovskii, *Theory of Plasma Instabilities*. Springer, 2013.

[135] G. C. Dyer, X. Shi, B. V. Olson, S. D. Hawkins, J. F. Klem, E. A. Shaner, and W. Pan, "Far infrared edge photoresponse and persistent edge transport in an inverted InAs/GaSb heterostructure," *Applied Physics Letters*, vol. 108, no. 1, p. 013106, Jan. 2016.

[136] J. Wu, A. S. Mayorov, C. D. Wood, D. Mistry, L. Li, W. Muchenje, M. C. Rosamond, L. Chen, E. H. Linfield, A. G. Davies, and J. E. Cunningham, "Excitation, detection, and electrostatic manipulation of terahertz-frequency range plasmons in a two-dimensional electron system," *Scientific Reports*, vol. 5, p. 15420, Oct. 2015.

[137] M. Dyakonov and M. Shur, "Detection, mixing, and frequency multiplication of terahertz radiation by two-dimensional electronic fluid," *IEEE Transactions on Electron Devices*, vol. 43, no. 3, pp. 380–387, Mar. 1996.

[138] V. V. Popov, "The resonant terahertz response of a slot diode with a two-dimensional electron channel," *Semiconductors*, vol. 39, no. 1, pp. 142–146, 2005.

[139] T. Otsuji, M. Hanabe, T. Nishimura, and E. Sano, "A grating-bicoupled plasma-wave photomixer with resonant-cavity enhanced structure," *Optics*

Express, vol. 14, no. 11, p. 4815, May 2006.

[140] M. Dyakonov and M. S. Shur, "Current instability and plasma waves generation in ungated two-dimensional electron layers," *Applied Physics Letters*, vol. 87, no. 11, p. 111501, Sep. 2005.

[141] D. Hofstetter, L. Diehl, J. Faist, W. J. Schaff, J. Hwang, L. F. Eastman, and C. Zellweger, "Midinfrared intersubband absorption on AlGaN/GaN-based high-electron-mobility transistors," *Applied Physics Letters*, vol. 80, no. 16, pp. 2991–2993, Apr. 2002.

[142] A. E. Fatimy, N. Dyakonova, Y. Meziani, T. Otsuji, W. Knap, S. Vandenbrouk, K. Madjour, D. Théron, C. Gaquiere, M. A. Poisson, S. Delage, P. Prystawko, and C. Skierbiszewski, "AlGaN/GaN high electron mobility transistors as a voltage-tunable room temperature terahertz sources," *Journal of Applied Physics*, vol. 107, no. 2, p. 024504, Jan. 2010.

[143] S. Rabbaa and J. Stiens, "Charge density and plasmon modes in a triangular quantum well model for doped and undoped gated AlGaN/GaN HEMTs," *Journal of Physics D: Applied Physics*, vol. 44, no. 32, p. 325103, Jul. 2011.

[144] M. Nakayama, "Theory of surface waves coupled to surface carriers," *Journal of the Physical Society of Japan*, vol. 36, no. 2, pp. 393–398, Feb. 1974.

[145] A. Eguiluz, T. K. Lee, J. J. Quinn, and K. W. Chiu, "Interface excitations in metal-insulator-semiconductor structures," *Physical Review B*, vol. 11, no. 12, pp. 4989–4993, Jun. 1975.

[146] V. V. Popov, A. N. Koudymov, M. Shur, and O. V. Polischuk, "Tuning of ungated plasmons by a gate in the field-effect transistor with two-dimensional

electron channel," *Journal of Applied Physics*, vol. 104, no. 2, p. 024508, Jul. 2008.

[147] V. V. Popov, M. S. Shur, G. M. Tsymbalov, and D. V. Fateev, "Higher-order plasmon reson in GaN-based field-effect transistor arrays," *International Journal of High Speed Electronics and Systems*, vol. 17, no. 03, pp. 557–566, Sep. 2007.

[148] Y. Zhao, N. Engheta, and A. Alù, "Homogenization of plasmonic metasurfaces modeled as transmission-line loads," *Metamaterials*, vol. 5, no. 2-3, pp. 90–96, Jun. 2011.

[149] A. Pors and S. I. Bozhevolnyi, "Plasmonic metasurfaces for efficient phase control in reflection," *Optics Express*, vol. 21, no. 22, p. 27438, Nov. 2013.